

GenCore version 5.1.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

February 16, 2006, 01:14:10 ; Search time 132 Seconds

(without alignments)
1115.090 Million cell updates/sec

Title: US-10-733-782-1

Perfect score: 1678

Sequence: 1 AIDENKQKALAAALGQIEKQ.....AKBIEKKVRELLSNPNSTP 335

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_21.*

- 1: Geneseqp1980a.*
- 2: Geneseqp1990a.*
- 3: Geneseqp2000a.*
- 4: Geneseqp2001a.*
- 5: Geneseqp2002a.*
- 6: Geneseqp2003a.*
- 7: Geneseqp2003bs.*
- 8: Geneseqp2004s.*
- 9: Geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1678	100.0	335	8 ADR16428	Adr16428 E. coli R
2	1678	100.0	352	8 ADR16432	Adr16432 E. coli w
3	1678	100.0	353	6 ABU15395	Abu15395 Protein e
4	1678	100.0	358	3 AAY68827	Aay68827 Amino aci
5	1678	100.0	358	3 AAY68829	Aay68829 Amino aci
6	1678	100.0	358	8 ADQ87824	Adq87824 E. coli r
7	1678	100.0	358	8 ADQ87822	Adq87822 E. coli r
8	1678	100.0	358	8 ADQ87821	Adq87821 E. coli r
9	1677	99.9	358	3 AAY68831	Aay68831 Amino aci
10	1674	99.8	335	8 ADR16430	Adr16430 E. coli R
11	1674	99.8	358	2 AAW64213	Aaw64213 New minsh
12	1673	99.7	358	3 AAY68828	Aay68828 Amino aci
13	1673	99.7	358	8 ADQ87823	Adq87823 E. coli r
14	1669	99.5	358	8 ADQ87826	Adq87826 E. coli r
15	1668	99.4	358	2 AAW64215	Aaw64215 Hyperreco
16	1665	99.2	358	2 AAW64214	Aaw64214 Hyperreco
17	1662	99.0	358	2 AAW64217	Aaw64217 Hyperreco
18	1662	99.0	358	2 AAW64218	Aaw64218 Hyperreco
19	1660	98.9	358	3 AAY68832	Aay68832 Amino aci
20	1660	98.9	358	8 ADQ87827	Adq87827 E. coli r
21	1659	98.9	358	3 AAY68830	Aay68830 Amino aci
22	1659	98.9	358	8 ADQ87825	Adq87825 E. coli r
23	1643	97.9	358	2 AAW64216	Aaw64216 Hyperreco
24	1641	97.8	353	6 ABU48024	Abu48024 Protein e

25	1613	96.1	352	6 ABU31601	Abu31601 Protein e
26	1613	96.1	356	7 ABO63433	ABO63433 Klebsiell
27	1593	94.9	352	6 ABU27708	Abu27708 Protein e
28	1547	92.2	356	6 ABM67793	ABM67793 Phototrab
29	1534	91.4	356	6 ABU50600	ABU50600 Protein e
30	1504	89.6	355	6 ABU40869	ABU40869 Protein e
31	1504	89.6	367	7 ADF05150	Adf05150 Bacteriol
32	1467	87.4	351	6 ABU44971	Abu44971 Protein e
33	1436	85.6	412	6 ABU49151	Abu49151 Protein e
34	1353	80.6	354	6 ABU39491	Abu39491 Protein e
35	1332	79.4	354	6 ABU30311	ABU30311 Protein e
36	1269	75.6	398	7 ABO72381	ABO72381 Pseudomon
37	1268	75.6	346	6 ABU38581	ABU38581 Protein e
38	1259	75.0	349	6 ABU16844	ABU16844 Protein e
39	1259	75.0	355	6 ADA36276	Ada36276 Acinetoba
40	1257	74.9	355	6 ABU40018	ABU40018 Protein e
41	1255	74.8	369	6 ABU41878	ABU41878 Protein e
42	1250.5	74.5	366	8 ADL05025	Adl05025 M. catarr
43	1245.5	74.2	349	6 ABU35468	ABU35468 Protein e
44	1241.5	74.0	348	6 ABU33125	ABU33125 Protein e
45	1241.5	74.0	348	9 AEB39711	Aeb39711 L. pneumo

ALIGNMENTS

RESULT 1

ADRI16428

ID ADRI16428 standard; protein; 335 AA.

XX ADRI16428;

DT 04-NOV-2004 (first entry)

DE E. coli RecA mutant protein #1.

XX RecA; DNA binding protein; ssDNA binding protein; SSB;

KW DNA strand exchange reaction; site-specific cleavage; mutant; muten.

XX Escherichia coli.

XX US2004157248-A1.

XX 12-AUG-2004.

PF 11-DEC-2003; 2003US-00733782.

XX 12-DEC-2002; 2002US-0432758P.

XX (COXM/) COX M M.

PA (LUSE/) LUSETTI S L.

PA (EGGL/) EGGLER A L.

XX (HARU/) HARUTA N.

XX Cox MM, Lusettti SL, Eggler AL, Haruta N;

XX WPI; 2004-580265/56.

XX N-PSDB; ADRI16429.

XX New RecA mutant proteins comprising a single mutation or a double mutation, useful for catalyzing homologous DNA pairing and DNA strand exchange reactions in an in vitro or in vivo environment.

PS Claim 2; SEQ ID NO 1; 34pp; English.

XX The invention relates to an isolated RecA mutant protein, where the protein is either a single mutant RecA protein comprising a deletion of amino acid residues from the carboxyl terminus, or a double mutant RecA protein comprising a deletion of amino acid residues from the carboxyl terminus and an amino acid change from a glutamate to a basic amino acid. The invention also relates to a polynucleotide sequence that encodes the RecA mutant protein, a method of catalyzing an in vitro homologous DNA pairing and DNA strand exchange reactions comprising providing an amount

CC of the RecA mutant protein and a method of increasing recombination
CC efficiency of homologous DNA pairing and DNA strand exchange reactions in
CC a cell comprising supplying to the cell an amount of the RecA mutant
CC protein. The RecA mutant protein comprises an enhanced capacity to
CC displace a DNA binding protein as compared to wild-type RecA. The DNA
CC binding protein is the ssDNA binding protein from *Escherichia coli* (SSB).
CC It also comprises enhanced binding to DNA during a DNA strand exchange
CC reaction as compared to wild-type RecA. The composition and methods are
CC useful for catalyzing homologous DNA pairing and DNA strand exchange
CC reactions in an in vitro or in vivo environment. These may be used in
CC promoting in vitro alterations of genes to permit the rapid construction
CC of desired gene mutants for industrial and pharmaceutical purposes. The
CC mutant proteins may also be used in targeting the site-specific cleavage
CC of small and large DNAs, or as a basis for the design and construction of
CC tiny electronic circuits based on DNA. This sequence represents a RecA
CC mutant protein of the invention.

XX Sequence 335 AA;

Query Match 100.0%; Score 1678; DB 8; Length 335;
Best Local Similarity 100.0%; Pred. No. 5.7e-155;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AIDENKQKALAAALGQIEKQFGKSGIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 60
DB 1 AIDENKQKALAAALGQIEKQFGKSGIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 60
QY 61 IVEIYGPSSSGKTTLTQVIAAAQREGKTCFIDEAHLDPYARKLVGVDIDNLLCSQPD 120
DB 61 IVEIYGPSSSGKTTLTQVIAAAQREGKTCFIDEAHLDPYARKLVGVDIDNLLCSQPD 120
QY 121 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGEIGDSHMGIAARMMSQAMRKLAG 180
DB 121 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGEIGDSHMGIAARMMSQAMRKLAG 180
QY 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVGEGENVVGS 240
DB 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVGEGENVVGS 240
QY 241 ETRVKVVKNKIAAPFKQAEFQILYGEINFGYELVDLGVKEKLEKAGAWYSYKGEKIQ 300
DB 241 ETRVKVVKNKIAAPFKQAEFQILYGEINFGYELVDLGVKEKLEKAGAWYSYKGEKIQ 300
QY 301 GRANATAWLKONPETAKEIEKKVRELLSNPNSTP 335
DB 301 GRANATAWLKONPETAKEIEKKVRELLSNPNSTP 335

RESULT 2
ID ADR16432
AC ADR16432 standard; protein; 352 AA.
XX ADR16432;
XX AC ADR16432;
XX 04-NOV-2004 (first entry)
XX E. coli wild-type RecA protein.
XX RecA; DNA binding protein; ssDNA binding protein; SSB;
KW DNA strand exchange reaction; site-specific cleavage.
XX *Escherichia coli*.
XX US2004157248-A1.
XX 12-AUG-2004.
XX 11-DEC-2003; 2003US-00733782.
XX 12-DEC-2002; 2002US-0432758P.
XX (COXM/) COX M M.
PA (LUSE/) LUSSETTI S L.
PA

PA (EGGL/) EGGLER A L.
PA (HARU/) HARUTA N.
XX Cox MM, Lusetti SL, Egglar AL, Haruta N;
XX WPI; 2004-580265/56.
XX New RecA mutant proteins comprising a single mutation or a double
PT mutation, useful for catalyzing homologous DNA pairing and DNA strand
PT exchange reactions in an in vitro or in vivo environment.
XX Disclosure; SEQ ID NO 5; 34pp; English.
XX The invention relates to an isolated RecA mutant protein, where the
CC protein is either a single mutant RecA protein comprising a deletion of
CC amino acid residues from the carboxyl terminus, or a double mutant RecA
CC protein comprising a deletion of amino acid residues from the carboxyl
CC terminus and an amino acid change from a glutamate to a basic amino acid.
CC The invention also relates to a polynucleotide sequence that encodes the
CC RecA mutant protein, a method of catalyzing an in vitro homologous DNA
CC pairing and DNA strand exchange reactions comprising providing an amount
CC of the RecA mutant protein and a method of increasing recombination
CC efficiency of homologous DNA pairing and DNA strand exchange reactions in
CC a cell comprising supplying to the cell an amount of the RecA mutant
CC protein. The RecA mutant protein comprises an enhanced capacity to
CC displace a DNA binding protein as compared to wild-type RecA. The DNA
CC binding protein is the ssDNA binding protein from *Escherichia coli* (SSB).
CC It also comprises enhanced binding to DNA during a DNA strand exchange
CC reaction as compared to wild-type RecA. The composition and methods are
CC useful for catalyzing homologous DNA pairing and DNA strand exchange
CC reactions in an in vitro or in vivo environment. These may be used in
CC promoting in vitro alterations of genes to permit the rapid construction
CC of desired gene mutants for industrial and pharmaceutical purposes. The
CC mutant proteins may also be used in targeting the site-specific cleavage
CC of small and large DNAs, or as a basis for the design and construction of
CC tiny electronic circuits based on DNA. This sequence represents the wild-
XX type RecA protein of the invention.

SQ Sequence 352 AA;

Query Match 100.0%; Score 1678; DB 8; Length 352;
Best Local Similarity 100.0%; Pred. No. 6.2e-155;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AIDENKQKALAAALGQIEKQFGKSGIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 60
DB 1 AIDENKQKALAAALGQIEKQFGKSGIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 60
QY 61 IVEIYGPSSSGKTTLTQVIAAAQREGKTCFIDEAHLDPYARKLVGVDIDNLLCSQPD 120
DB 61 IVEIYGPSSSGKTTLTQVIAAAQREGKTCFIDEAHLDPYARKLVGVDIDNLLCSQPD 120
QY 121 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGEIGDSHMGIAARMMSQAMRKLAG 180
DB 121 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGEIGDSHMGIAARMMSQAMRKLAG 180
QY 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVGEGENVVGS 240
DB 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVGEGENVVGS 240
QY 241 ETRVKVVKNKIAAPFKQAEFQILYGEINFGYELVDLGVKEKLEKAGAWYSYKGEKIQ 300
DB 241 ETRVKVVKNKIAAPFKQAEFQILYGEINFGYELVDLGVKEKLEKAGAWYSYKGEKIQ 300
QY 301 GRANATAWLKONPETAKEIEKKVRELLSNPNSTP 335
DB 301 GRANATAWLKONPETAKEIEKKVRELLSNPNSTP 335

RESULT 3
ID ABU15395
XX ABU15395 standard; protein; 353 AA.
XX

AC ABUL5395;
XX 19-JUN-2003 (first entry)
XX Protein encoded by Prokaryotic essential gene #922.
XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX Escherichia coli.
XX WO200277183-A2.
XX 03-OCT-2002.
XX 21-MAR-2002; 2002WO-US009107.
XX 21-MAR-2001; 2001US-00815242.
XX 06-SEP-2001; 2001US-00948993.
XX 25-OCT-2001; 2001US-0342923P.
XX 08-FEB-2002; 2002US-00072851.
XX 06-MAR-2002; 2002US-0362699P.
XX (ELIT-) ELITRA PHARM INC.
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-029926/02.
XX N-PSDB; ACA19265.
XX New antisense nucleic acids, useful for identifying proteins or screening
XX for homologous nucleic acids required for cellular proliferation to
XX isolate candidate molecules for rational drug discovery programs.
XX Claim 25; SEQ ID NO 43319; 1766pp; English.
XX The invention relates to an isolated nucleic acid comprising any one of
XX the 6213 antisense sequences given in the specification where expression
XX of the nucleic acid inhibits proliferation of a cell. Also included are:
XX (1) a vector comprising a promoter operably linked to the nucleic acid
XX encoding a polypeptide whose expression is inhibited by the antisense
XX nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX polypeptide or its fragment whose expression is inhibited by the
XX antisense nucleic acid; (4) an antibody capable of specifically binding
XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX proliferation or the activity of a gene in an operon required for
XX proliferation; (7) identifying a compound that influences the activity of
XX the gene product or that has an activity against a biological pathway
XX required for proliferation, or that inhibits cellular proliferation; (8)
XX identifying a gene required for cellular proliferation or the biological
XX pathway in which a proliferation-required gene or its gene product lies
XX or a gene on which the test compound that inhibits proliferation of an
XX organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX compound's activity; (11) a culture comprising strains in which the gene
XX product is overexpressed or underexpressed; (12) determining the extent
XX to which each of the strains is present in a culture or collection of
XX strains; or (13) identifying the target of a compound that inhibits the
XX proliferation of an organism. The antisense nucleic acids are useful for
XX identifying proteins or screening for homologous nucleic acids required
XX for cellular proliferation to isolate candidate molecules for rational
XX drug discovery programs, or for screening homologous nucleic acids
XX required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
XX *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
XX the target prokaryotic essential genes. Note: The sequence data for this
XX patent did not form part of the printed specification, but was obtained
XX in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 353 AA;
Query Match 100.0%; Score 1678; DB 6; Length 353;
Best Local Similarity 100.0%; Pred. No. 6.2e-155;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AIDENKOKALAAALQIQIEKQFGKGSIMRLGEDRSMDVETISTGSLSLDIALCAGGLPMGR 60
DB |||||
QY 2 AIDENKOKALAAALQIQIEKQFGKGSIMRLGEDRSMDVETISTGSLSLDIALCAGGLPMGR 61
DB |||||
QY 61 IVEIYGPSSSGKTTLTQVIAAAQREGTKCAPIDAEHALDPIYARKLGVDIDNLLCSQPD 120
DB |||||
QY 62 IVEIYGPSSSGKTTLTQVIAAAQREGTKCAPIDAEHALDPIYARKLGVDIDNLLCSQPD 121
DB |||||
QY 121 TGEQALEICDALARSGAVDVIVDSVAALTPKAEITEGIGSHMGLAARMMSOAMRKLKAG 180
DB |||||
QY 122 TGEQALEICDALARSGAVDVIVDSVAALTPKAEITEGIGSHMGLAARMMSOAMRKLKAG 181
DB |||||
QY 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGNNALKFYASVRLDRIIRRGAVKEGNNVGS 240
DB |||||
QY 182 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGNNALKFYASVRLDRIIRRGAVKEGNNVGS 241
DB |||||
QY 241 ETRVKVVKNTIAAPFKQAEFQILYGEINFGYELVDLGKVKLEKAGAWSYKGEKIGQ 300
DB |||||
QY 242 ETRVKVVKNTIAAPFKQAEFQILYGEINFGYELVDLGKVKLEKAGAWSYKGEKIGQ 301
DB |||||
QY 301 GKANATAWLKDNPETAKEIEKKVRELLLSNPNSTP 335
DB |||||
QY 302 GKANATAWLKDNPETAKEIEKKVRELLLSNPNSTP 336
DB |||||

RESULT 4

AY68827
ID AY68827 standard; protein; 358 AA.
XX AC AY68827;
XX DT 16-MAY-2000 (first entry)
XX DE Amino acid sequence of a wildtype recA protein.
XX KW recA; hyperrecombinogenic variant; male gamete; female gamete;
KW heat tolerance; ethanol production; ethanol tolerance; metabolite; taxol;
KW cyclosporin A; erythromycin; meiosis.
XX OS Escherichia coli.
XX FN WO200004190-A1.
XX PD 27-JAN-2000.
XX PF 15-JUL-1999; 99WO-US015972.
XX PR 15-JUL-1998; 98US-00116188.
XX PA (MAXY-) MAXYGEN INC.
XX PI Del Cardayre S, Tobin M, Stemmer WPC, Ness JE, Minshull J;
PI Patten PA, Subramanian V, Castle LA, Krebber CM, Bass S, Zhang Y;
PI Cox T, Huiegan G, Yuan L, Affholter JA;
DR WPI: 2000-182446/16.
DR N-PSDB; AAZ60609.
XX PT Evolution of whole cells and organism by iterative cycles of
XX recombination and selection and screening for acquisition of desired
XX properties.
XX PS Example 1; Fig 13; 197pp; English.
XX CC The present sequence represents the recA protein. The wild type recA
XX protein was used to produce hyperrecombinogenic variants (see AAY68828-
XX 32), using the method of the invention. The specification describes a
XX method for producing a library of diverse multicellular organisms using
XX pools of male and female gametes. At least one of the male pool or female
XX pools comprises a number of different gametes derived from different
XX strains of a species or of a different species. The viable organisms
XX produced from the fertilized gametes are repeatedly crossed to produce a

CC library of diverse organisms, which are selected for a desired trait or
CC property. The methods can be used to evolve cells to acquire a desired
CC property such as heat tolerance, ethanol production or tolerance, acid,
CC improved production and maintenance of enzyme cofactors or NAD(P)H and
CC improved glucose transport. The desired property may be expression of a
CC protein or primary or secondary metabolite. Alternatively the desired
CC property is secretion of a protein or secondary metabolite, chosen from
CC taxol, cyclosporin A and erythromycin. The desired property may be a
CC capacity for meiosis or compatibility to form a heterokaryon with another
CC strain
XX
SQ Sequence 358 AA;

Query Match 100.0%; Score 1678; DB 3; Length 358;
Best Local Similarity 100.0%; Pred. No. 6.3e-155;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLDLALGAGGLPMGR 60
DB 7 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLDLALGAGGLPMGR 66
QY 61 IVEIYGPESSGKTTLTQVIAAAQREGKTCFIDEHALDPIYARKLGVDIDNLLCSQPD 120
DB 67 IVEIYGPESSGKTTLTQVIAAAQREGKTCFIDEHALDPIYARKLGVDIDNLLCSQPD 126
QY 121 TGEQALEICDALARSGAVDVIIVDSVAALTTPKAEIEGEIGDSHMGLAARMMSQAMRKLAG 180
DB 127 TGEQALEICDALARSGAVDVIIVDSVAALTTPKAEIEGEIGDSHMGLAARMMSQAMRKLAG 186
QY 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVG 240
DB 187 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVG 246
QY 241 ETRVKVVKNKIAAPFKQAEFQILYGEINFGYELVDLGVEKLEKAGAWYSYKGEKIQ 300
DB 247 ETRVKVVKNKIAAPFKQAEFQILYGEINFGYELVDLGVEKLEKAGAWYSYKGEKIQ 306
QY 301 GKANATAMLKONPETAKETIEKKVRELLSNPNSTP 335
DB 307 GKANATAMLKONPETAKETIEKKVRELLSNPNSTP 341

RESULT 5
AAY68829
ID AAY68829 standard; protein; 358 AA.
XX
AC AAY68829;
XX
DT 16-MAY-2000 (first entry)
XX
DE Amino acid sequence of a hyperrecombinogenic recA protein clone 4.
XX
KW recA; hyperrecombinogenic variant; male gamete; female gamete;
KW heat tolerance; ethanol production; ethanol tolerance; metabolite; taxol;
KW cyclosporin A; erythromycin; meiosis.
XX
OS Synthetic.
OS Escherichia coli.
XX
PN W0200004190-A1.
XX
PD 27-JAN-2000.
XX
PF 15-JUL-1999; 99WO-US015972.
XX
PR 15-JUL-1999; 98US-00116188.
XX
PA (MAXY-) MAXYGEN INC.
XX
PI Del Cardayre S, Tobin M, Stemmer WPC, Ness JE, Minshull J;
PI Patten PA, Subramanian V, Castle LA, Kriebler CM, Baas S, Zhang Y;
PI Cox T, Huiean G, Yuan L, Affholter JA;
XX

DR WPI; 2000-182446/16.
DR N-PSDB; AAZ60611.
XX
PT Evolution of whole cells and organism by iterative cycles of
PT recombination and selection and screening for acquisition of desired
PT properties.
XX
PS Example 1; Fig 13; 197pp; English.
XX
CC The present sequence represents a hyperrecombinogenic recA protein. The
CC wild type recA protein (AAY68827) was used to produce hyperrecombinogenic
CC variants, using the method of the invention. The specification describes
CC a method for producing a library of diverse multicellular organisms using
CC pools of male and female gametes. At least one of the male pool or female
CC pools comprises a number of different gametes. The viable organisms
CC strains of a species or of a different species. The desired property is
CC produced from the fertilized gametes are repeatedly crossed to produce a
CC library of diverse organisms, which are selected for a desired trait or
CC property. The methods can be used to evolve cells to acquire a desired
CC property such as heat tolerance, ethanol production or tolerance, acid,
CC improved production and maintenance of enzyme cofactors or NAD(P)H and
CC improved glucose transport. The desired property may be expression of a
CC protein or primary or secondary metabolite. Alternatively the desired
CC property is secretion of a protein or secondary metabolite, chosen from
CC taxol, cyclosporin A and erythromycin. The desired property may be a
CC capacity for meiosis or compatibility to form a heterokaryon with another
CC strain
XX
SQ Sequence 358 AA;

Query Match 100.0%; Score 1678; DB 3; Length 358;
Best Local Similarity 100.0%; Pred. No. 6.3e-155;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLDLALGAGGLPMGR 60
DB 7 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLDLALGAGGLPMGR 66
QY 61 IVEIYGPESSGKTTLTQVIAAAQREGKTCFIDEHALDPIYARKLGVDIDNLLCSQPD 120
DB 67 IVEIYGPESSGKTTLTQVIAAAQREGKTCFIDEHALDPIYARKLGVDIDNLLCSQPD 126
QY 121 TGEQALEICDALARSGAVDVIIVDSVAALTTPKAEIEGEIGDSHMGLAARMMSQAMRKLAG 180
DB 127 TGEQALEICDALARSGAVDVIIVDSVAALTTPKAEIEGEIGDSHMGLAARMMSQAMRKLAG 186
QY 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVG 240
DB 187 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVG 246
QY 241 ETRVKVVKNKIAAPFKQAEFQILYGEINFGYELVDLGVEKLEKAGAWYSYKGEKIQ 300
DB 247 ETRVKVVKNKIAAPFKQAEFQILYGEINFGYELVDLGVEKLEKAGAWYSYKGEKIQ 306
QY 301 GKANATAMLKONPETAKETIEKKVRELLSNPNSTP 335
DB 307 GKANATAMLKONPETAKETIEKKVRELLSNPNSTP 341

RESULT 6
ADQ87824
ID ADQ87824 standard; protein; 358 AA.
XX
AC ADQ87824;
XX
DT 09-SEP-2004 (first entry)
XX
XX E. coli recA protein clone #3.
XX
KW RecA; protoplast formation; gamete; heat tolerance; ethanol production;
KW ethanol tolerance; enzyme cofactor; NAD(P)H; glucose transport; meiosis;
KW heterokaryon; taxol; cyclosporin A; erythromycin; viral infection;
KW plant genome shuffling; microprobe manipulation; reiterative pooling.

XX OS Escherichia coli.
 XX AU2004200501-A1.
 XX PD 04-MAR-2004.
 XX PF 09-FEB-2004; 2004AU-00200501.
 XX PR 09-FEB-2004; 2004AU-00200501.
 XX PA (MAXY-) MAXYGEN INC.
 XX PI Yuan L, Cox T, Bass S, Castle LA, Patten PA, Ness JE, Tobin M;
 PI Affholter JA, Huismann G, Zhang Y, Krebber CM, Subramanian V;
 PI Minshull J, Stemmer WPC, Del Cardayre S;
 DR WPI; 2004-507924/49.
 DR N-PSDB; ADQ87817.
 XX

Evolving cells to acquire a desired property, by forming protoplasts of different cells, fusing protoplasts to form hybrid protoplasts (HP), producing regenerated cells, forming additional HP and producing additional regenerated cells.

PS Disclosure; Fig 13; 196pp; English.

XX The invention relates to a method of evolving cells to acquire a desired property by forming protoplasts of different cells, fusing the protoplasts to form hybrid protoplasts, incubating the hybrid protoplasts to produce regenerated cells, repeatedly forming protoplasts from regenerated cells, fusing protoplasts to form hybrid protoplasts in which genomes from protoplasts recombine to form additional hybrid genomes and incubating additional hybrid protoplasts for producing additional regenerated cells. The invention also relates to a method of producing a library of diverse multicellular organisms involving providing a pool of male gametes and a pool of female gametes, where one of the male pool or the female pool comprises several different gametes derived from different strains of a species or different species and the male gametes fertilise the female gametes, permitting at least a portion of the resulting fertilised gametes to grow into reproductively viable organisms, repeatedly crossing the reproductively viable organisms to produce a library of diverse organisms and selecting the library for a desired trait or property. The methods are useful for evolving cells to acquire desired properties such as heat tolerance, ethanol production, ethanol tolerance, improved production and maintenance of enzyme cofactors, improved production and maintenance of NAD(P)H and improved glucose transport. The desired property is the expression of a protein, primary metabolite or secondary metabolite, the secretion of a protein or heterokaryon with another strain. The secondary metabolite is chosen from taxol, cyclosporin A and erythromycin. The method is also useful for acquisition of desired properties such as enhanced recombination, gene copy number or gene reductivity and capacity for expression and/or secretion of proteins or secondary metabolites. The method is further useful for predicting efficacy of a drug in treating viral infection, plant genome shuffling, micropore manipulation, producing transgenic animals, improvement of overexpressed genes for a desired phenotype and reiterative pooling and breeding of higher organisms. The cells evolved by the method are useful in molecular genetics. This sequence represents an E. coli recA protein clone used in the method of the invention.

SQ Sequence 358 AA;

Query Match 100.08; Score 1678; DB 8; Length 358;
 Best Local Similarity 100.08; Pred. No. 6.3e-155;
 Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AIDENKQKALAAALGQIEKQFGKSGIMRLGDRSMVDVETISTGSLSDIALGAGLPMGR 60
 DB 7 AIDENKQKALAAALGQIEKQFGKSGIMRLGDRSMVDVETISTGSLSDIALGAGLPMGR 66
 QY 61 IVEITYGPSSGKTTTLTQVIAAAQREGKTCAFIDAEHALDPIYARKLGVLDNLLCSQPD 120

Db 67 IVEITYGPSSGKTTTLTQVIAAAQREGKTCAFIDAEHALDPIYARKLGVLDNLLCSQPD 126
 QY 121 TGEQALEICDALARSGAVDVIVVDSVAALTPKAEIEGEIGDSHMGIAARMMSQAMRKLKAG 180
 Db 127 TGEQALEICDALARSGAVDVIVVDSVAALTPKAEIEGEIGDSHMGIAARMMSQAMRKLKAG 186
 QY 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVUGS 240
 Db 187 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVUGS 246
 QY 241 ETRVKVVKNKTAAPFKQAEFQILYOGGINFYGELVDLGVKEKLIKAGAWYSYKGEKIQG 300
 Db 247 ETRVKVVKNKTAAPFKQAEFQILYOGGINFYGELVDLGVKEKLIKAGAWYSYKGEKIQG 306
 QY 301 GKANATWLKONPETAKEIEKKVRELLLSNPNSTP 335
 Db 307 GKANATWLKONPETAKEIEKKVRELLLSNPNSTP 341

RESULT 7
ADQ87822

ID ADQ87822 standard; protein; 358 AA.

XX AC ADQ87822;

XX DT 09-SEP-2004 (first entry)

XX DE E. coli recA protein clone #1.

XX KW RecA; protoplast formation; gamete; heat tolerance; ethanol production;
 KW ethanol tolerance; enzyme cofactor; NAD(P)H; glucose transport; meiosis;
 KW heterokaryon; taxol; cyclosporin A; erythromycin; viral infection;
 KW plant genome shuffling; micropore manipulation; reiterative pooling.

XX OS Escherichia coli.

XX AU2004200501-A1.

XX PD 04-MAR-2004.

XX PF 09-FEB-2004; 2004AU-00200501.

XX PR 09-FEB-2004; 2004AU-00200501.

XX PA (MAXY-) MAXYGEN INC.

XX PI Yuan L, Cox T, Bass S, Castle LA, Patten PA, Ness JE, Tobin M;
 PI Affholter JA, Huismann G, Zhang Y, Krebber CM, Subramanian V;
 PI Minshull J, Stemmer WPC, Del Cardayre S;

XX DR WPI; 2004-507924/49.

XX DR N-PSDB; ADQ87815.

XX Evolving cells to acquire a desired property, by forming protoplasts of
 XX different cells, fusing protoplasts to form hybrid protoplasts (HP),
 XX producing regenerated cells, forming additional HP and producing
 XX additional regenerated cells.

XX PS Disclosure; Fig 13; 196pp; English.

XX The invention relates to a method of evolving cells to acquire a desired
 XX property by forming protoplasts of different cells, fusing the
 XX protoplasts to form hybrid protoplasts, incubating the hybrid protoplasts
 XX to produce regenerated cells, repeatedly forming protoplasts from
 XX regenerated cells, fusing protoplasts to form hybrid protoplasts in which
 XX genomes from protoplasts recombine to form additional hybrid genomes and
 XX incubating additional hybrid protoplasts for producing additional
 XX regenerated cells. The invention also relates to a method of producing a
 XX library of diverse multicellular organisms involving providing a pool of
 XX male gametes and a pool of female gametes, where one of the male pool or
 XX the female pool comprises several different gametes derived from
 XX different strains of a species or different species and the male gametes

CC fertilise the female gametes, permitting at least a portion of the
CC resulting fertilised gametes to grow into reproductively viable
CC organisms, repeatedly crossing the reproductively viable organisms to
CC produce a library of diverse organisms and selecting the library for a
CC desired trait or property. The methods are useful for evolving cells to
CC acquire desired properties such as heat tolerance, ethanol production,
CC ethanol tolerance, improved production and maintenance of enzyme
CC cofactors, improved production and maintenance of NAD(P)H and improved
CC glucose transport. The desired property is the expression of a protein, or
CC primary metabolite or secondary metabolite, the secretion of a protein or
CC secondary metabolite, capacity for meiosis or compatibility to form a
CC heterokaryon with another strain. The secondary metabolite is chosen from
CC taxol, cyclosporin A and erythromycin. The method is also useful for
CC acquisition of desired properties such as enhanced recombination, gene
CC expression of proteins or secondary metabolites. The method is further
CC useful for predicting efficacy of a drug in treating viral infection.
CC plant genome shuffling, micropore manipulation, producing transgenic
CC animals, improvement of overexpressed genes for a desired phenotype and
CC reiterative pooling and breeding of higher organisms. The cells evolved
CC by the method are useful in molecular genetics. This sequence represents
CC an E. coli recA protein clone used in the method of the invention.
XX
XX Sequence 358 AA;

Query Match 100.0%; Score 1678; DB 8; Length 358;
Best Local Similarity 100.0%; Pred. No. 6.3e-155;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSDMDVETITGSLDIALGAGGLPMGR 60
DB 7 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSDMDVETITGSLDIALGAGGLPMGR 66
QY 61 IVEIYGPSSGKTTTLQVIAAAQREGKTCAFIDAEHALDPIYARKLGVDDIDNLLCSQPD 120
DB 67 IVEIYGPSSGKTTTLQVIAAAQREGKTCAFIDAEHALDPIYARKLGVDDIDNLLCSQPD 126
QY 121 TGEQALEICDALARSGAVDVIIVVDSVAALTPKAEIEGETGDSHMGIAARMMSQAMRKLKAG 180
DB 127 TGEQALEICDALARSGAVDVIIVVDSVAALTPKAEIEGETGDSHMGIAARMMSQAMRKLKAG 186
QY 181 NLKQSTLLIFINQIRMKIIVGMPGNPTTTGGNALKFYASVRLDIRIGAVGEGENVVGS 240
DB 187 NLKQSTLLIFINQIRMKIIVGMPGNPTTTGGNALKFYASVRLDIRIGAVGEGENVVGS 246
QY 241 ETRVKVVKNKIAAPFKQAEFQILYGGEGINFGYELVDLGVKEKLIKAGAWYSYKGEKIQ 300
DB 247 ETRVKVVKNKIAAPFKQAEFQILYGGEGINFGYELVDLGVKEKLIKAGAWYSYKGEKIQ 306
QY 301 GKANATAWLKNDNPETAKEIEKKVRELLSNPNSTP 335
DB 307 GKANATAWLKNDNPETAKEIEKKVRELLSNPNSTP 341

RESULT 8
ADQ87821
ID ADQ87821 standard; protein; 358 AA.
AC ADQ87821;
XX
XX 09-SEP-2004 (first entry)
DT
XX
XX E. coli recA protein.
XX
XX RecA; protoplast formation; gamete; heat tolerance; ethanol production;
KW ethanol tolerance; enzyme cofactor; NAD(P)H; glucose transport; meiosis;
KW heterokaryon; taxol; cyclosporin A; erythromycin; viral infection;
KW plant genome shuffling; micropore manipulation; reiterative pooling.
XX Escherichia coli.
XX
XX AU2004200501-A1.
PN
XX

PD 04-MAR-2004.
XX 09-FEB-2004; 2004AU-00200501.
XX 09-FEB-2004; 2004AU-00200501.
XX (MAXY-) MAXYGEN INC.
XX Yuan L, Cox T, Bass S, Castle LA, Patten PA, Ness JE, Tobin M;
PI Affholter JA, Huismann G, Zhang Y, Kriebler CW, Subramanian V;
PI Minshull J, Stemmer WPC, Del Cardayre S;
XX WPI: 2004-507924/49.
DR N-PSDB; ADQ87814.
XX
XX Evolving cells to acquire a desired property, by forming protoplasts of
PT different cells, fusing protoplasts to form hybrid protoplasts (HP),
PT producing regenerated cells, forming additional HP and producing
PT additional regenerated cells.
XX
XX Disclosure; Fig 13; 196pp; English.
XX
XX The invention relates to a method of evolving cells to acquire a desired
CC property by forming protoplasts of different cells, fusing the
CC protoplasts to form hybrid protoplasts, incubating the hybrid protoplasts
CC to produce regenerated cells, repeatedly forming protoplasts from
CC regenerated cells, fusing protoplasts to form hybrid protoplasts in which
CC genomes from protoplasts recombine to form additional hybrid genomes and
CC incubating additional hybrid protoplasts for producing additional
CC regenerated cells. The invention also relates to a method of producing a
CC library of diverse multicellular organisms involving providing a pool of
CC male gametes and a pool of female gametes, where one of the male pool or
CC the female pool comprises several different gametes derived from
CC different strains of a species or different species and the male gametes
CC fertilise the female gametes, permitting at least a portion of the
CC resulting fertilised gametes to grow into reproductively viable
CC organisms, repeatedly crossing the reproductively viable organisms to
CC produce a library of diverse organisms and selecting the library for a
CC desired trait or property. The methods are useful for evolving cells to
CC acquire desired properties such as heat tolerance, ethanol production,
CC ethanol tolerance, improved production and maintenance of enzyme
CC cofactors, improved production and maintenance of NAD(P)H and improved
CC glucose transport. The desired property is the expression of a protein,
CC primary metabolite or secondary metabolite, the secretion of a protein or
CC secondary metabolite, capacity for meiosis or compatibility to form a
CC heterokaryon with another strain. The secondary metabolite is chosen from
CC taxol, cyclosporin A and erythromycin. The method is also useful for
CC acquisition of desired properties such as enhanced recombination, gene
CC expression of proteins or secondary metabolites. The method is further
CC useful for predicting efficacy of a drug in treating viral infection,
CC plant genome shuffling, micropore manipulation, producing transgenic
CC animals, improvement of overexpressed genes for a desired phenotype and
CC reiterative pooling and breeding of higher organisms. The cells evolved
CC by the method are useful in molecular genetics. This sequence represents
CC the E. coli recA protein used in the method of the invention.
XX
XX Sequence 358 AA;

Query Match 100.0%; Score 1678; DB 8; Length 358;
Best Local Similarity 100.0%; Pred. No. 6.3e-155;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSDMDVETITGSLDIALGAGGLPMGR 60
DB 7 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSDMDVETITGSLDIALGAGGLPMGR 66
QY 61 IVEIYGPSSGKTTTLQVIAAAQREGKTCAFIDAEHALDPIYARKLGVDDIDNLLCSQPD 120
DB 67 IVEIYGPSSGKTTTLQVIAAAQREGKTCAFIDAEHALDPIYARKLGVDDIDNLLCSQPD 126
QY 121 TGEQALEICDALARSGAVDVIIVVDSVAALTPKAEIEGETGDSHMGIAARMMSQAMRKLKAG 180

Db 127 TGEQALEICDALARGANDVIVVDSVAALTPKAEIEGEIGDSHMGSLAARMMSQAMKLAG 186
 QY 181 NLKQSNLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVG 240
 Db 187 NLKQSNLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVG 246
 QY 241 ETRVKVVKNKIAAPFQKAEFQILYGGINFGYBELVGLVKEKLEIKAGAWSYKGEKIQ 300
 Db 247 ETRVKVVKNKIAAPFQKAEFQILYGGINFGYBELVGLVKEKLEIKAGAWSYKGEKIQ 306
 QY 301 GKANATAWLKNDPETAKEIEKKVRELLSNPNSTP 335
 Db 307 GKANATAWLKNDPETAKEIEKKVRELLSNPNSTP 341

RESULT 9

AAV68831
 ID AAV68831 standard; protein; 358 AA.

XX AAV68831;

DT 16-MAY-2000 (first entry)

DE Amino acid sequence of a hyperrecombinogenic recA protein clone 6.

XX recA; hyperrecombinogenic variant; male gamete; female gamete;
 KW heat tolerance; ethanol production; ethanol tolerance; metabolite; taxol;
 KW cyclosporin A; erythromycin; meiosis.

OS Synthetic.
 OS Escherichia coli.

XX Key Location/Qualifiers

FT Misc-difference 156

FT /note= "encoded by TCG"

PN WO200004190-A1.

XX 27-JAN-2000.

PF 15-JUL-1999; 99WO-US015972.

XX 15-JUL-1998; 98US-00116188.

XX (MAXY-) MAXYGEN INC.

PI Del Cardayre S, Tobin M, Stemmer WPC, Ness JE, Minshull J; Zhang Y;
 PI Patten PA, Subramanian V, Casale LA, Krebber CM, Bass S, Zhang Y;
 PI Cox T, Huisman G, Yuan L, Alfholter JA;

XX WPI; 2000-182446/16.

DR N-PSDB; AA260613.

XX Evolution of whole cells and organism by iterative cycles of
 PT recombination and selection and screening for acquisition of desired
 PT properties.

XX Example 1; Fig 13; 197pp; English.

XX The present sequence represents a hyperrecombinogenic recA protein. The
 CC wild type recA protein (AAV68827) was used to produce hyperrecombinogenic
 CC variants, using the method of the invention. The specification describes
 CC a method for producing a library of diverse multicellular organisms using
 CC pools of male and female gametes. At least one of the male pool or female
 CC pools comprises a number of different gametes derived from different
 CC strains of a species or of a different species. The viable organisms
 CC produced from the fertilized gametes are repeatedly crossed to produce a
 CC library of diverse organisms, which are selected for a desired trait or
 CC property. The methods can be used to evolve cells to acquire a desired
 CC property such as heat tolerance, ethanol production or tolerance, acid,
 CC improved production and maintenance of enzyme cofactors or NAD(P)H and
 CC improved glucose transport. The desired property may be expression of a
 CC protein or primary or secondary metabolite. Alternatively the desired

CC property is secretion of a protein or secondary metabolite, chosen from
 CC taxol, cyclosporin A and erythromycin. The desired property may be a
 CC capacity for meiosis or compatibility to form a heterokaryon with another
 CC strain
 XX Sequence 358 AA;

Query Match 99.9%; Score 1677; DB 3; Length 358;
 Best Local Similarity 99.7%; Pred. No. 7.9e-155;
 Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AIDENKQKALAAALGQIEKQFGKSGIMRLGEDRSDMDVETISTGSLDLALGAGGLPMGR 60
 Db 7 AIDENKQKALAAALGQIEKQFGKSGIMRLGEDRSDMDVETISTGSLDLALGAGGLPMGR 66

QY 61 IVEIYGPSSSGKTTTLQVIAAQAQREGKTCAPIDAHAHALDPIYARKLGVDDINLLCSQPD 120
 Db 67 IVEIYGPSSSGKTTTLQVIAAQAQREGKTCAPIDAHAHALDPIYARKLGVDDINLLCSQPD 126

QY 121 TGEQALEICDALARGANDVIVVDSVAALTPKAEIEGEIGDSHMGSLAARMMSQAMKLAG 180
 Db 127 TGEQALEICDALARGANDVIVVDSVAALTPKAEIEGEIGDSHMGSLAARMMSQAMKLAG 186

QY 181 NLKQSNLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVG 240
 Db 187 NLKQSNLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVG 246

QY 241 ETRVKVVKNKIAAPFQKAEFQILYGGINFGYBELVGLVKEKLEIKAGAWSYKGEKIQ 300
 Db 247 ETRVKVVKNKIAAPFQKAEFQILYGGINFGYBELVGLVKEKLEIKAGAWSYKGEKIQ 306

QY 301 GKANATAWLKNDPETAKEIEKKVRELLSNPNSTP 335
 Db 307 GKANATAWLKNDPETAKEIEKKVRELLSNPNSTP 341

RESULT 10

ADRI6430

ID ADRI6430 standard; protein; 335 AA.

XX ADRI6430;

DT 04-NOV-2004 (first entry)

DE E. coli RecA mutant protein #2.

KW RecA; DNA binding protein; ssDNA binding protein; SSB;
 KW DNA strand exchange reaction; site-specific cleavage; mutant; muten.

OS Escherichia coli.

XX Key Location/Qualifiers

FT Misc-difference 38

FT /note= "Wild-type Glu substituted by Lys"

PN US2004157248-A1.

XX 12-AUG-2004.

PF 11-DEC-2003; 2003US-00733782.

XX 12-DEC-2002; 2002US-0432758P.

XX (COXM/) COX M M.

PA (LUSE/) LUSETTI S L.

PA (EGGL/) EGGLER A L.

PA (HARU/) HARUTA N.

PI Cox MM, Lusettti SL, Eggler AL, Haruta N;

XX WPI; 2004-580265/56.

DR N-PSDB; ADRI6431.

XX

PT New RecA mutant proteins comprising a single mutation or a double
PT mutation, useful for catalyzing homologous DNA pairing and DNA strand
PT exchange reactions in an in vitro or in vivo environment.
XX
PS Claim 13; SEQ ID NO 3; 34pp; English.
XX
CC The invention relates to an isolated RecA mutant protein, where the
CC protein is either a single mutant RecA protein comprising a deletion of
CC amino acid residues from the carboxyl terminus, or a double mutant RecA
CC protein comprising a deletion of amino acid residues from the carboxyl
CC terminus and an amino acid change from a glutamate to a basic amino acid.
CC The invention also relates to a polynucleotide sequence that encodes the
CC RecA mutant protein, a method of catalysing an in vitro homologous DNA
CC pairing and DNA strand exchange reactions comprising providing an amount
CC of the RecA mutant protein and a method of increasing recombination
CC efficiency of homologous DNA pairing and DNA strand exchange reactions in
CC a cell comprising supplying to the cell an amount of the RecA mutant
CC protein. The RecA mutant protein comprises an enhanced capacity to
CC displace a DNA binding protein as compared to wild-type RecA. The DNA
CC binding protein is the ssDNA binding protein from Escherichia coli (SSB).
CC It also comprises enhanced binding to DNA during a DNA strand exchange
CC reaction as compared to wild-type RecA. The composition and methods are
CC useful for catalyzing homologous DNA pairing and DNA strand exchange
CC reactions in an in vitro or in vivo environment. These may be used in
CC promoting in vitro alterations of genes to permit the rapid construction
CC of desired gene mutants for industrial and pharmaceutical purposes. The
CC mutant proteins may also be used in targeting the site-specific cleavage
CC of small and large DNAs, or as a basis for the design and construction of
CC tiny electronic circuits based on DNA. This sequence represents a RecA
CC mutant protein of the invention.
XX
SQ Sequence 335 AA;

Query Match 99.8%; Score 1674; DB 8; Length 335;
Best Local Similarity 99.7%; Pred. No. 1.4e-154;
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 AIDENKQKALAAALGQIEKQFGKSGIMRLGEDRSMVDVETISTGSLDLIALGAGGLPMGR 60
DB 1 AIDENKQKALAAALGQIEKQFGKSGIMRLGEDRSMVDVETISTGSLDLIALGAGGLPMGR 60
QY 61 IVEIYGPSSGKTTLTQVIAAAQREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
DB 61 IVEIYGPSSGKTTLTQVIAAAQREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
QY 121 TGEQALEICDALARSGAVDIVVDSVAALTTPKAEIEGEIGDSHMGLAARMWSQAMRKL 180
DB 121 TGEQALEICDALARSGAVDIVVDSVAALTTPKAEIEGEIGDSHMGLAARMWSQAMRKL 180
QY 181 NLKQSNLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVG 240
DB 181 NLKQSNLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVG 240
QY 241 ETRVKVVKNKIAAPFKQAEFQILYGEINFGYELVDLGVKKEKLEKAGAWSYKGEKIG 300
DB 241 ETRVKVVKNKIAAPFKQAEFQILYGEINFGYELVDLGVKKEKLEKAGAWSYKGEKIG 300
QY 301 GKANATAWLKDNPETAKEIEKKVRELLLSNPSTP 335
DB 301 GKANATAWLKDNPETAKEIEKKVRELLLSNPSTP 335

RESULT 11
AAW64213
ID AAW64213 standard; protein; 358 AA.
XX AAW64213;
XX
DT 28-APR-1999 (first entry)
XX
DE New minshall recA protein.
XX
KW Escherichia coli; recA protein; hyperrecombinogenic variant; Minshall;

KW recursive sequence recombination; evolution.
XX Escherichia coli.
XX WO9831837-A1.
XX
XX 23-JUL-1998.
XX
XX 16-JAN-1998; 98WO-US0000852.
XX
XX 17-JAN-1997; 97US-0035054P.
XX
XX (MAXY-) MAXYGEN INC.
XX
XX Delcardayre SB, Tobin MB, Stemmer WPC, Ness JE, Minshall J;
PI Patten P;
XX
XX WPI; 1998-427565/60.
XX
XX N-PSDB; AA44285.
XX
XX Evolution of whole cells and organisms by recursive DNA sequence
XX recombination in cells to evolve cells having acquired desired function,
XX useful in methods for predicting the efficacy of a drug in treating viral
XX or pathogenic infections.
XX
XX Example 1; Fig 13; 125pp; English.
XX
XX The present invention provides methods employing iterative cycles of
XX recombination and selection/screening for evolution of whole cells and
XX organisms toward acquisition of desired properties, e.g. enhanced
XX recombinogenicity, genome copy number, and capacity for expression and/or
XX secretion of proteins and secondary metabolites. The present sequence
XX represents a wild-type recA protein (designated new Minshall), from an
XX example of the present invention
XX
SQ Sequence 358 AA;

Query Match 99.8%; Score 1674; DB 2; Length 358;
Best Local Similarity 99.7%; Pred. No. 1.6e-154;
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 AIDENKQKALAAALGQIEKQFGKSGIMRLGEDRSMVDVETISTGSLDLIALGAGGLPMGR 60
DB 7 AIDENKQKALAAALGQIEKQFGKSGIMRLGEDRSMVDVETISTGSLDLIALGAGGLPMGR 66
QY 61 IVEIYGPSSGKTTLTQVIAAAQREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
DB 67 IVEIYGPSSGKTTLTQVIAAAQREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 126
QY 121 TGEQALEICDALARSGAVDIVVDSVAALTTPKAEIEGEIGDSHMGLAARMWSQAMRKL 180
DB 127 TGEQALEICDALARSGAVDIVVDSVAALTTPKAEIEGEIGDSHMGLAARMWSQAMRKL 186
QY 181 NLKQSNLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVG 240
DB 187 NLKQSNLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVG 246
QY 241 ETRVKVVKNKIAAPFKQAEFQILYGEINFGYELVDLGVKKEKLEKAGAWSYKGEKIG 300
DB 247 ETRVKVVKNKIAAPFKQAEFQILYGEINFGYELVDLGVKKEKLEKAGAWSYKGEKIG 306
QY 301 GKANATAWLKDNPETAKEIEKKVRELLLSNPSTP 335
DB 307 GKANATAWLKDNPETAKEIEKKVRELLLSNPSTP 341

RESULT 12
AAV68828
ID AAV68828 standard; protein; 358 AA.
XX
XX AAV68828;
XX
DT 16-MAY-2000 (first entry)

XX Amino acid sequence of a hyperrecombinogenic recA protein clone 2.
 DE recA; hyperrecombinogenic variant; male gamete; female gamete;
 KW heat tolerance; ethanol production; ethanol tolerance; metabolite; taxol;
 KW cyclosporin A; erythromycin; meiosis.
 XX Synthetic.
 OS Escherichia coli.
 XX WO200004190-A1.
 XX 27-JAN-2000.
 PD
 XX 15-JUL-1999; 99WO-US015972.
 PF
 XX 15-JUL-1998; 98US-00116188.
 PR
 XX (MAXY-) MAXYGEN INC.
 PA
 XX Del Cardayre S, Tobin M, Stemmer WPC, Ness JE, Minshull J;
 PI Patten PA, Subramanian V, Castle LA, Krebber CM, Bass S, Zhang Y;
 PI Cox T, Huisman G, Yuan L, Affholter JA;
 XX
 DR WPI; 2000-182446/16.
 DR N-PSDB; AAZ60610.
 XX
 XX Evolution of whole cells and organism by iterative cycles of
 PT recombination and selection and screening for acquisition of desired
 PT properties.
 XX
 PS Example 1; Fig 13; 197pp; English.
 XX
 CC The present sequence represents a hyperrecombinogenic recA protein. The
 CC wild type recA protein (AAV68827) was used to produce hyperrecombinogenic
 CC variants, using the method of the invention. The specification describes
 CC a method for producing a library of diverse multicellular organisms using
 CC pools of male and female gametes. At least one of the male pool or female
 CC pools comprises a number of different gametes derived from different
 CC strains of a species or of a different species. The viable organisms
 CC produced from the fertilized gametes are repeatedly crossed to produce a
 CC library of diverse organisms, which are selected for a desired trait or
 CC property. The methods can be used to evolve cells to acquire a desired
 CC property such as heat tolerance, ethanol production or tolerance, acid,
 CC improved production and maintenance of enzyme cofactors or NAD(P)H and
 CC improved glucose transport. The desired property may be expression of a
 CC protein or primary or secondary metabolite. Alternatively the desired
 CC property is secretion of a protein or secondary metabolite, chosen from
 CC taxol, cyclosporin A and erythromycin. The desired property may be a
 CC capacity for meiosis or compatibility to form a heterokaryon with another
 CC strain
 XX
 SQ Sequence 358 AA;
 Query Match 99.7%; Score 1673; DB 3; Length 358;
 Best Local Similarity 99.4%; Pred. No. 1.9e-154;
 Matches 333; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 AIDENKQALAAALGQIEKQFGKSGIMRLGEDRSMVETISTGSLDLALGAGLPMGR 60
 DB 7 AIDENKQALATGALGQIEKQFGKSGIMRLGEDRSMVETISTGSLDLALGAGLPMGR 66
 QY 61 IVEIYGPESSGKTTTLQVIAAAREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
 DB 67 IVEIYGPESSGKTTTLQVIAAAREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 126
 QY 121 TGEQALEICDALARGVDVIVDSVAALTPKAEIEGEGISGHMGLAARMSQARKLAG 180
 DB 127 TGEQALEICDALARGVDVIVDSVAALTPKAEIEGEGISGHMGLAARMSQARKLAG 186
 QY 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGNAKLFYASVRLDIRRIGAVKEGENVVGVS 240
 DB 187 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGNAKLFYASVRLDIRRIGAVKEGENVVGVS 246

QY 241 ETRVVKVKNKIAAPFKQAEFQILYGEINFGYELVDLGVKELIEKAGAWYSYKGEKIQG 300
 DB 247 ETRVVKVKNKIAAPFKQAEFQILYGEINFGYELVDLGVKELIEKAGAWYSYKGEKIQG 306
 QY 301 GHANATAWLKONPETAKIEKKVRELLLSNPNSTP 335
 DB 307 GHANATAWLKONPETAKIEKKVRELLLSNPNSTP 341

RESULT 13 ADQ87823

ID ADQ87823 standard; protein; 358 AA.

XX AC ADQ87823;

DT 09-SEP-2004 (first entry)

XX DE E. coli recA protein clone #2.

XX KW recA; protoplast formation; gamete; heat tolerance; ethanol production;
 KW ethanol tolerance; enzyme cofactor; NAD(P)H; glucose transport; meiosis;
 KW heterokaryon; taxol; cyclosporin A; erythromycin; viral infection;
 KW plant genome shuffling; micropore manipulation; reiterative pooling.

XX OS Escherichia coli.

FN AU2004200501-A1.

XX PD 04-MAR-2004.

XX PF 09-FEB-2004; 2004AU-00200501.

XX PR 09-FEB-2004; 2004AU-00200501.

XX PA (MAXY-) MAXYGEN INC.

PI Yuan L, Cox T, Bass S, Castle LA, Patten PA, Ness JE, Tobin M;
 PI Affholter JA, Huisman G, Zhang Y, Krebber CM, Subramanian V;
 PI Minshull J, Stemmer WPC, Del Cardayre S;

XX WPI; 2004-507924/49.

DR N-PSDB; ADQ87816.

XX Evolving cells to acquire a desired property, by forming protoplasts of
 PT different cells, fusing protoplasts to form hybrid protoplasts (HP),
 PT producing regenerated cells, forming additional HP and producing
 PT additional regenerated cells.

PS Disclosure; Fig 13; 196pp; English.

XX The invention relates to a method of evolving cells to acquire a desired
 CC property by forming protoplasts of different cells, fusing the
 CC protoplasts to form hybrid protoplasts, incubating the hybrid protoplasts
 CC to produce regenerated cells, repeatedly forming protoplasts from
 CC regenerated cells, fusing protoplasts to form hybrid protoplasts in which
 CC genomes from protoplasts recombine to form additional hybrid genomes and
 CC incubating additional hybrid protoplasts for producing additional
 CC regenerated cells. The invention also relates to a method of producing a
 CC library of diverse multicellular organisms involving providing a pool of
 CC male gametes and a pool of female gametes, where one of the male pool or
 CC the female pool comprises several different gametes derived from
 CC different strains of a species or different species and the male gametes
 CC fertilise the female gametes, permitting at least a portion of the
 CC resulting fertilised gametes to grow into reproductively viable
 CC organisms, repeatedly crossing the reproductively viable organisms to
 CC produce a library of diverse organisms and selecting the library for a
 CC desired trait or property. The methods are useful for evolving cells to
 CC acquire desired properties such as heat tolerance, ethanol production,
 CC ethanol tolerance, improved production and maintenance of enzyme
 CC cofactors, improved production and maintenance of NAD(P)H and improved
 CC glucose transport. The desired property is the expression of a protein,
 CC primary metabolite or secondary metabolite, the secretion of a protein or

CC secondary metabolite, capacity for meiosis or compatibility to form a
CC heterokaryon with another strain. The secondary metabolite is chosen from
CC taxol, cyclosporin A and erythromycin. The method is also useful for
CC acquisition of desired properties such as enhanced recombination, or
CC gene copy number or gene reductivity and capacity for expression and/or
CC secretion of proteins or secondary metabolites. The method is further
CC useful for predicting efficacy of a drug in treating viral infection,
CC plant genome shuffling, micropore manipulation, producing transgenic
CC animals, improvement of overexpressed genes for a desired phenotype and
CC reiterative pooling and breeding of higher organisms. The cells evolved
CC by the method are useful in molecular genetics. This sequence represents
CC an E. coli recA protein clone used in the method of the invention.

XX Sequence 358 AA;

Query Match 99.7%; Score 1673; DB 8; Length 358;
Best Local Similarity 99.4%; Pred. No. 1.9e-154;
Matches 333; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 AIDENKQKALAAALGQIEKQFGKSGIMRLGEDRSMVDVETISTGSLDLALGAGGLPMGR 60
DB 7 AIDENKQKALAAALGQIEKQFGKSGIMRLGEDRSMVDVETISTGSLDLALGAGGLPMGR 66
QY 61 IVEIYGPESSGKTTTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
DB 67 IVEIYGPESSGKTTTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 126
QY 121 TGEQALEICDALARSGANDVIVVDSVAALTTPKAEIEGEIGDSHGLAARMMSQAMKLAG 180
DB 127 TGEQALEICDALARSGANDVIVVDSVAALTTPKAEIEGEIGDSHGLAARMMSQAMKLAG 186
QY 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVGEGENVVGS 240
DB 187 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVGEGENVVGS 246
QY 241 ETRVKVVKNKIAAPFKQAEFQILYGEINFGYELVDLGKVEKLEKAGAWSYKGEKIQ 300
DB 247 ETRVKVVKNKIAAPFKQAEFQILYGEINFGYELVDLGKVEKLEKAGAWSYKGEKIQ 306
QY 301 GRANATAWLKNDPETAKEKKVRELLSNPNSTP 335
DB 307 GRANATAWLKNDPETAKEKKVRELLSNPNSTP 341

RESULT 14
ADQ87826
ID ADQ87826 standard; protein; 358 AA.

XX ADQ87826;

XX 09-SEP-2004 (first entry)

XX E. coli recA protein clone #5.

XX RecA; protoplast formation; gamete; heat tolerance; ethanol production;
XX ethanol tolerance; enzyme cofactor; NAD(P)H; glucose transport; meiosis;
XX heterokaryon; taxol; cyclosporin A; erythromycin; viral infection;
XX plant genome shuffling; micropore manipulation; reiterative pooling.

OS Escherichia coli.

XX AU2004200501-A1.

XX 04-MAR-2004.

XX 09-FEB-2004; 2004AU-00200501.

XX 09-FEB-2004; 2004AU-00200501.

XX (MAXY-) MAXYGEN INC.

XX Yuan L, Cox T, Bass S, Castle LA, Patten PA, Ness JE, Tobin M;
XX Affholter JA, Huisman G, Zhang Y, Krebber CM, Subramanian V;

PI

Minshull J, Stemmer WPC, Del Cardayre S;

WPI; 2004-507924/49.

N-PSDB; ADQ87819.

Evolving cells to acquire a desired property, by forming protoplasts of
different cells, fusing protoplasts to form hybrid protoplasts (HP),
producing regenerated cells, forming additional HP and producing
additional regenerated cells.

Dislosure; Fig 13; 196pp; English.

The invention relates to a method of evolving cells to acquire a desired
property by forming protoplasts of different cells, fusing the
protoplasts to form hybrid protoplasts, incubating the hybrid protoplasts
to produce regenerated cells, repeatedly forming protoplasts from
regenerated cells, fusing protoplasts to form hybrid protoplasts in which
genomes from protoplasts recombine to form additional hybrid genomes and
incubating additional hybrid protoplasts for producing additional
regenerated cells. The invention also relates to a method of producing a
library of diverse multicellular organisms involving providing a pool of
male gametes and a pool of female gametes, where one of the male pool or
the female pool comprises several different gametes derived from
different strains of a species or different species and the male gametes
fertilise the female gametes, permitting at least a portion of the
resulting fertilised gametes to grow into reproductively viable
organisms, repeatedly crossing the reproductively viable organisms to
produce a library of diverse organisms and selecting the library for a
desired trait or property. The methods are useful for evolving cells to
acquire desired properties such as heat tolerance, ethanol production,
ethanol tolerance, improved production and maintenance of enzyme
cofactors, improved production and maintenance of NAD(P)H and improved
glucose transport. The desired property is the expression of a protein,
primary metabolite or secondary metabolite, the secretion of a protein or
heterokaryon with another strain. The secondary metabolite is chosen from
taxol, cyclosporin A and erythromycin. The method is also useful for
acquisition of desired properties such as enhanced recombination, or
gene copy number or gene reductivity and capacity for expression and/or
secretion of proteins or secondary metabolites. The method is further
useful for predicting efficacy of a drug in treating viral infection,
plant genome shuffling, micropore manipulation, producing transgenic
animals, improvement of overexpressed genes for a desired phenotype and
reiterative pooling and breeding of higher organisms. The cells evolved
by the method are useful in molecular genetics. This sequence represents
an E. coli recA protein clone used in the method of the invention.

XX Sequence 358 AA;

Query Match 99.5%; Score 1669; DB 8; Length 358;

Best Local Similarity 99.4%; Pred. No. 4.8e-154;

Matches 333; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AIDENKQKALAAALGQIEKQFGKSGIMRLGEDRSMVDVETISTGSLDLALGAGGLPMGR 60

DB 7 AIDENKQKALAAALGQIEKQFGKSGIMRLGEDRSMVDVETISTGSLDLALGAGGLPMGR 66

QY 61 IVEIYGPESSGKTTTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 120

DB 67 IVEIYGPESSGKTTTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 126

QY 121 TGEQALEICDALARSGANDVIVVDSVAALTTPKAEIEGEIGDSHGLAARMMSQAMKLAG 180

DB 127 TGEQALEICDALARSGANDVIVVDSVAALTTPKAEIEGEIGDSHGLAARMMSQAMKLAG 186

QY 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVGEGENVVGS 240

DB 187 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVGEGENVVGS 246

QY 241 ETRVKVVKNKIAAPFKQAEFQILYGEINFGYELVDLGKVEKLEKAGAWSYKGEKIQ 300

DB 247 ETRVKVVKNKIAAPFKQAEFQILYGEINFGYELVDLGKVEKLEKAGAWSYKGEKIQ 306

QY 301 GKANATAWLKONPETAKIEKKVRELLLSNPSTP 335
Db 307 GKANATAWLKONPETAKIEKKVRELLLSNPSTP 341

RESULT 15

AAW64215
ID AAW64215 standard; protein; 358 AA.

XX AC AAW64215;
XX 28-APR-1999 (first entry)
XX Hyperrecombinogenic variant recA protein clone 4.

XX Escherichia coli; recA protein; hyperrecombinogenic variant; Minshall;
KW recursive sequence recombination; evolution.
XX Escherichia coli.

OS Synthetic.
XX WO9831837-A1.

PN 23-JUL-1998.

PD 16-JAN-1998; 98WO-US000852.

PF 17-JAN-1997; 97US-0035054P.

PR (MAXY-) MAXYGEN INC.

XX Delcardayre SB, Tobin MB, Stemmer WPC, Ness JE, Minshall J;
PI Patten P;
XX WPI: 1998-427565/60.
DR N-PSDB; AAV44287.

XX Evolution of whole cells and organisms by recursive DNA sequence
PT recombination in cells to evolve cells having acquired desired function,
PT useful in methods for predicting the efficacy of a drug in treating viral
PT or pathogenic infections.

XX Example 1; Fig 13; 125pp; English.

XX The present invention provides methods employing iterative cycles of
CC recombination and selection/screening for evolution of whole cells and
CC organisms toward acquisition of desired properties, e.g. enhanced
CC recombination, genome copy number, and capacity for expression and/or
CC secretion of proteins and secondary metabolites. The present sequence
CC represents a hyperrecombinogenic variant of a recA protein (see
CC AAW64213), from an example of the present invention

XX Sequence 358 AA;

Query Match 99.4%; Score 1668; DB 2; Length 358;
Best Local Similarity 99.1%; Pred. No. 6e-154;
Matches 332; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 AIDENKQKALAAALQIEKQFGKSGIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 60
Db 7 AIDENKQKALAAALQIEKQFGKSGIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 66
QY 61 IVEIYGPSGGKTTLTQVIAAQRGKTCAPIDAEHALDPIYARKGVDDNLLCSQPD 120
Db 67 IVEIYGPSGGKTTLTQVIAAQRGKTCAPIDAEHALDPIYARKGVDDNLLCSQPD 126
QY 121 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGIDSHMGLAARMMSQAMRKLKAG 180
Db 127 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGIDSHMGLAARMMSQAMRKLKAG 186
QY 181 NLKQNTLLIFINQIRMKIGVFNPGPPTTGGNALKFYASVRLDIRRIGAVKEGENVYGS 240
Db 187 NLKQNTLLIFINQIRMKIGVFNPGPPTTGGNALKFYASVRLDIRRIGAVKEGENVYGS 246

QY 241 ETRVKVVKNKIAAPFKQAEFQILYGEINFGYELVDLGVKEKLIBKAGAWYSYKGEKIQG 300
Db 247 ETRVKVVKNKIAAPFKQAEFQILYGEINFGYELVDLGVKEKLIBKAGAWYSYKGEKIQG 306
QY 301 GKANATAWLKONPETAKIEKKVRELLLSNPSTP 335
Db 307 GKANATAWLKONPETAKIEKKVRELLLSNPSTP 341

Search completed: February 16, 2006, 01:18:51
Job time : 135 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: February 16, 2006, 01:19:09 ; Search time 25.5 Seconds
(without alignments)
1264.024 Million cell updates/sec

Title: US-10-733-782-1
Perfect score: 1678
Sequence: 1 AIDENKQKALAAALGQIEKQ.....AKETIKVRELLSNPNSTP 335
Scoring table: BLOSUN62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1678	100.0	353	1 ROECA	recombination prot
2	1678	100.0	353	2 D91073	RecA protein [impo
3	1666	99.3	353	2 H85917	hypothetical prote
4	1641	97.8	353	2 AG0843	RecA protein [impo
5	1534	91.4	356	2 AG0401	RecA protein [impo
6	1532	91.3	354	2 S31481	recombination prot
7	1530	91.2	356	2 S37586	recombination prot
8	1504	89.6	355	1 ROEBPM	recombination prot
9	1436	85.6	412	2 E82310	recA protein VC054
10	1417	84.4	354	2 S46274	recombination prot
11	1412.5	84.2	354	2 JQ1461	recombination prot
12	1332	79.4	348	2 A49929	recombination prot
13	1323	78.8	353	2 JCS198	recombination prot
14	1268	75.6	346	1 RQPSAA	RecA protein PA361
15	1265	75.4	349	2 JN0321	recombination prot
16	1264.5	75.4	344	2 JQ0776	recombination prot
17	1257	74.9	349	2 I39509	recombination prot
18	1255	74.8	355	2 T10482	recA protein - Pse
19	1250	74.5	352	2 JQ2406	recombination prot
20	1246	74.3	352	2 A47709	recombination prot
21	1238.5	73.8	348	1 A60989	recombination prot
22	1237.5	73.2	342	2 S08728	recombination prot
23	1217.5	72.6	348	2 F81082	RecA protein NMB14
24	1216.5	72.5	348	2 F81860	RecA protein NMA16
25	1213	72.3	378	2 A83350	recA protein [impo
26	1203.5	71.7	348	2 JQ0745	recombination prot
27	1203.5	71.7	351	2 S16896	recombination prot
28	1201.5	71.6	347	1 RQPSAC	recombination prot
29	1200.5	71.5	363	2 JQ1377	recombination prot

ALIGNMENTS

RESULT 1

ROECA

recombination protein recA [validated] - Escherichia coli (strain K-12)

N:Alternate names: recombinase A

C:Species: Escherichia coli

C>Date: 31-Jul-1980 #sequence revision 14-Nov-1997 #text change 09-Jul-2004

C:Accession: G65049; A93847; A93846; S11931; S63525; S69129; S63979; A03548

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;

.A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of Escherichia coli K-12.

A:Reference number: A64720; MUID:97426617; PMID:9278503

A:Accession: G65049

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-353 <BLAT>

A:Cross-references: UNIPROT:P03017; UNIPARC:UPI0000112C1C; GB:AE000354; GB:U00096; NID

R:Sancar, A.; Stachelek, C.; Konigsberg, W.; Rupp, W.D.

Proc. Natl. Acad. Sci. U.S.A. 77, 2611-2615, 1980

A:Title: Sequences of the recA gene and protein.

A:Reference number: A93846; MUID:80234673; PMID:6930655

A:Accession: A93846

A:Molecule type: DNA

A:Residues: 2-353 <HOR>

A:Cross-references: UNIPARC:UPI0000112C17; GB:V00328; NID:942672; PIDN:CAA23618.1; PID

R:Sancar, A.; Stachelek, C.; Konigsberg, W.; Rupp, W.D.

Proc. Natl. Acad. Sci. U.S.A. 77, 2611-2615, 1980

A:Title: Sequences of the recA gene and protein.

A:Reference number: A93846; MUID:80234673; PMID:6930655

A:Accession: A93846

A:Molecule type: DNA

A:Residues: 2-353 <SAN>

A:Cross-references: UNIPARC:UPI0000112C17; GB:V00328; NID:942672; PIDN:CAA23618.1; PID

R:Zhao, X.J.; McEntee, K.

Mol. Gen. Genet. 222, 369-376, 1990

A:Title: DNA sequence analysis of the recA genes from Proteus vulgaris, Erwinia carotov

A:Reference number: S11931; MUID:91109725; PMID:2274037

A:Accession: S11931

A:Molecule type: DNA

A:Residues: 2-353 <ZHA>

A:Cross-references: UNIPARC:UPI0000112C17; GB:X55552; NID:942678

A>Note: this ORF is not annotated in GenBank entry ECRECAGEN, release 109.0

R:Morimatsu, K.; Horii, T.

Eur. J. Biochem. 234, 695-705, 1995

A:Title: DNA-binding surface of RecA protein. Photochemical cross-linking of the first

A:Reference number: S63525; MUID:96096752; PMID:8529655

A:Accession: S63525

A:Status: preliminary

A:Molecule type: protein

A:Residues: 65-69;90-97;179-184;200-207;258-265;304-311;323-331 <MORI>

A:Cross-references: UNIPARC:UPI00001748A7; UNIPARC:UPI00001748A8; UNIPARC:UPI00001748A9

R;Morimatsu, K.; Horii, T.
Eur. J. Biochem. 228, 772-778, 1995
A:Title: The DNA-binding site of the RecA protein. Photochemical cross-linking of Tyr103
A:Reference number: S69129; MUID:95255284; PMID:7737176
A:Accession: S69129
A:Status: preliminary
A:Molecule type: protein
A:Residues: 90-108,180-184 <MOR2>
A:Cross-references: UNIPARC:UPI00001748AE
R;Gardner, R.V.; Voloshin, O.N.; Camerini-Otero, R.D.
Eur. J. Biochem. 233, 419-425, 1995
A:Title: The identification of the single-stranded DNA-binding domain of the Escherichia
A:Reference number: S63979; MUID:96067680; PMID:7588783
A:Accession: S63979
A:Molecule type: protein
A:Residues: 'XX',187-190,192-194 <GAR>
A:Cross-references: UNIPARC:UPI00001748AF
R;Yu, X.; Egelman, E.H.
submitted to the Brookhaven Protein Data Bank, December 1996
A:Reference number: A67277; PDB:2REC
A:Contents: annotation; X-ray crystallography, 2.3 angstroms, residues 4-157;166-195;211
R;Aihara, H.; Ito, Y.; Kurumizaka, H.; Terada, T.; Yokoyama, S.; Shibata, T.
submitted to the Brookhaven Protein Data Bank, January 1997
A:Reference number: A67455; PDB:1AA3
A:Contents: annotation; conformation by (1)H- and (15)N-NMR, residues 269-331
C:Genetics:
A:Gene: recA
A:Map position: 58 min
C:Function:
A:Description: plays an essential role in homologous recombination, in induction of the
C:Superfamily: recombination protein recA
C:Keywords: ATP; DNA binding; DNA recombination; DNA repair; nucleotide binding; P-loop;
F:67-74/Region: nucleotide-binding motif A (P-loop)
F:141-146/Region: nucleotide-binding motif B
F:73/Binding site: ATP (Lys) #status predicted

Query Match 100.0%; Score 1678; DB 1; Length 353;
Best Local Similarity 100.0%; Pred. No. 4.6e-106; Indels 0; Gaps 0;
Matches 335; Conservative 0; Mismatches 0

Qy 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 60
Db 2 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 61

Qy 61 IVEIYGPSSSGKTTTLQVIAAAQREGKTCAPIDAEHALDPIYARKLGVVDINLLCSQPD 120
Db 62 IVEIYGPSSSGKTTTLQVIAAAQREGKTCAPIDAEHALDPIYARKLGVVDINLLCSQPD 121

Qy 121 TGEQALEICDALARSGAVDVIVDSVAALTPKAEIEGIGDSHMGCLAARMMSQAMRKLKAG 180
Db 122 TGEQALEICDALARSGAVDVIVDSVAALTPKAEIEGIGDSHMGCLAARMMSQAMRKLKAG 181

Qy 181 NLKQSNLTLLIFINQIRMKIGVFNPGPTTTGNNALKFYASVRLDIRRICAVKEGENVVGS 240
Db 182 NLKQSNLTLLIFINQIRMKIGVFNPGPTTTGNNALKFYASVRLDIRRICAVKEGENVVGS 241

Qy 241 ETRVKVKNKTAAPFKQAEFQILYEGEINFGELVDLGVKEKLEKAGAWTSYKGEKIGQ 300
Db 242 ETRVKVKNKTAAPFKQAEFQILYEGEINFGELVDLGVKEKLEKAGAWTSYKGEKIGQ 301

Qy 301 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 335
Db 302 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 336

RESULT 2
D91073
RecA protein [imported] - Escherichia coli (strain O157:H7, substrain R1MD 0509952)
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: D91073
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: D91073
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-353 <HAY>
A:Cross-references: UNIPROT:P03017; UNIPARC:UPI0000112C1C; GB:BA000007; PIDN:BA036979.1,
A:Experimental source: strain O157:H7, substrain R1MD 0509952
C:Genetics:
A:Gene: EC63556
C:Superfamily: recombination protein recA

Query Match 100.0%; Score 1678; DB 2; Length 353;
Best Local Similarity 100.0%; Pred. No. 4.6e-106; Indels 0; Gaps 0;
Matches 335; Conservative 0; Mismatches 0

Qy 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 60
Db 2 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 61

Qy 61 IVEIYGPSSSGKTTTLQVIAAAQREGKTCAPIDAEHALDPIYARKLGVVDINLLCSQPD 120
Db 62 IVEIYGPSSSGKTTTLQVIAAAQREGKTCAPIDAEHALDPIYARKLGVVDINLLCSQPD 121

Qy 121 TGEQALEICDALARSGAVDVIVDSVAALTPKAEIEGIGDSHMGCLAARMMSQAMRKLKAG 180
Db 122 TGEQALEICDALARSGAVDVIVDSVAALTPKAEIEGIGDSHMGCLAARMMSQAMRKLKAG 181

Qy 181 NLKQSNLTLLIFINQIRMKIGVFNPGPTTTGNNALKFYASVRLDIRRICAVKEGENVVGS 240
Db 182 NLKQSNLTLLIFINQIRMKIGVFNPGPTTTGNNALKFYASVRLDIRRICAVKEGENVVGS 241

Qy 241 ETRVKVKNKTAAPFKQAEFQILYEGEINFGELVDLGVKEKLEKAGAWTSYKGEKIGQ 300
Db 242 ETRVKVKNKTAAPFKQAEFQILYEGEINFGELVDLGVKEKLEKAGAWTSYKGEKIGQ 301

Qy 301 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 335
Db 302 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 336

RESULT 3
H85917
hypothetical protein recA [imported] - Escherichia coli (strain O157:H7, substrain EDL933
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: H85917
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grobbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: H85917
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-353 <STO>
A:Cross-references: UNIPARC:UPI00001658E3; GB:AB005174; NID:gl2517139; PIDN:AAG57804.1;
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: recA
C:Superfamily: recombination protein recA

Query Match 99.3%; Score 1666; DB 2; Length 353;
Best Local Similarity 99.4%; Pred. No. 3e-105;
Matches 333; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 60
Db 2 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 61

Qy 61 IVEIYGPSSSGKTTTLQVIAAAQREGKTCAPIDAEHALDPIYARKLGVVDINLLCSQPD 120

Db 62 IVEIYGPSSGKTTTLTQVIAAQRGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 121
Qy 121 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGIDSHGMLGAARMSQAMRKLAG 180
Db 122 TGEQALEICDALARSGAVNVIIVDSVAALTPKAEIEXEIGDSHGMLGAARMSQAMRKLAG 181
Qy 181 NLKQSNLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVVG 240
Db 182 NLKQSNLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVVG 241
Qy 241 ETRVKVVKNKIAAPFKQAEFOLLYGEGINFYGVLDLGVKEKLEKAGAWTSYKGEKIGQ 300
Db 242 ETRVKVVKNKIAAPFKQAEFOLLYGEGINFYGVLDLGVKEKLEKAGAWTSYKGEKIGQ 301
Qy 301 GKANATAWLKONPETAKEIEKKVRELLLSNPNSTP 335
Db 302 GKANATAWLKONPETAKEIEKKVRELLLSNPNSTP 336

RESULT 4
AH0843
ReCA protein (imported) - Salmonella enterica subsp. enterica serovar Typhi (strain CT18)
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AH0843
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Park, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AH0843
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-353 <PAR>
A:Cross-references: UNIPARC:UPI000005A334; GB:AL513382; PIDN:CAD05935.1; PID:g16503906;
C:Genetics:
C:Superfamily: recombination protein recA

Query Match 97.8%; Score 1641; DB 2; Length 353;
Best Local Similarity 97.3%; Pred. No. 1.4e-103;
Matches 326; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
Qy 1 AIDENKOKAALAAALGOIEKQFGKGSIMRLGEDRSMVETITSGSLDIALGAGGLPMGR 60
Db 2 AIDENKOKAALAAALGOIEKQFGKGSIMRLGEDRSMVETITSGSLDIALGAGGLPMGR 61
Qy 61 IVEIYGPSSGKTTTLTQVIAAQRGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 62 IVEIYGPSSGKTTTLTQVIAAQRGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 121
Qy 121 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGIDSHGMLGAARMSQAMRKLAG 180
Db 122 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGIDSHGMLGAARMSQAMRKLAG 181
Qy 181 NLKQSNLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVVG 240
Db 182 NLKQSNLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVVG 241
Qy 241 ETRVKVVKNKIAAPFKQAEFOLLYGEGINFYGVLDLGVKEKLEKAGAWTSYKGEKIGQ 300
Db 242 ETRVKVVKNKIAAPFKQAEFOLLYGEGINFYGVLDLGVKEKLEKAGAWTSYKGEKIGQ 301
Qy 301 GKANATAWLKONPETAKEIEKKVRELLLSNPNSTP 335
Db 302 GKANATAWLKONPETAKEIEKKVRELLLSNPNSTP 336

RESULT 5
AG0401

RecA protein (imported) - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C:Accession: AG0401
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.;
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
Li, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall, N.
Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AG0401
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-356 <KUR>
A:Cross-references: UNIPROT:P37858; UNIPARC:UPI0000165AF8; GB:AL590842; PIDN:CAC92539;
C:Genetics:
A:Gene: recA
C:Superfamily: recombination protein recA

Query Match 91.4%; Score 1534; DB 2; Length 356;
Best Local Similarity 91.9%; Pred. No. 2.4e-96;
Matches 305; Conservative 14; Mismatches 13; Indels 0; Gaps 0;
Qy 1 AIDENKOKAALAAALGOIEKQFGKGSIMRLGEDRSMVETITSGSLDIALGAGGLPMGR 60
Db 2 AIDENKOKAALAAALGOIEKQFGKGSIMRLGEDRSMVETITSGSLDIALGAGGLPMGR 61
Qy 61 IVEIYGPSSGKTTTLTQVIAAQRGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 62 IVEIYGPSSGKTTTLTQVIAAQRGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 121
Qy 121 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGIDSHGMLGAARMSQAMRKLAG 180
Db 122 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGIDSHGMLGAARMSQAMRKLAG 181
Qy 181 NLKQSNLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVVG 240
Db 182 NLKQSNLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVVG 241
Qy 241 ETRVKVVKNKIAAPFKQAEFOLLYGEGINFYGVLDLGVKEKLEKAGAWTSYKGEKIGQ 300
Db 242 ETRVKVVKNKIAAPFKQAEFOLLYGEGINFYGVLDLGVKEKLEKAGAWTSYKGEKIGQ 301
Qy 301 GKANATAWLKONPETAKEIEKKVRELLLSNPN 332
Db 302 GKANATAWLKONPETAKEIEKKVRELLLSNPN 333

RESULT 6
S31481
recombination protein recA - Enterobacter agglomerans
N:Alternate names: recombinase A
C:Species: Enterobacter agglomerans
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C:Accession: S31481
R:Rappold, C.S.J.; Klingmueller, W.
submitted to the EMBL Data Library, January 1993
A:Description: Cloning and sequencing of the recA gene from Enterobacter agglomerans 33
A:Reference number: S31480
A:Accession: S31481
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-354 <RAP>
A:Cross-references: UNIPROT:P33037; UNIPARC:UPI000016EC3A; GB:L03291; EMBL:Z19517; NID:
C:Genetics:
A:Gene: recA
C:Superfamily: recombination protein recA
C:Keywords: ATP; DNA binding; DNA recombination; DNA repair; nucleotide binding; P-loop
P:67-74/Region: nucleotide-binding motif A (P-loop)
P:141-146/Region: nucleotide-binding motif B
F:73/Binding site: ATP (Lys) #status predicted
Query Match 91.3%; Score 1532; DB 2; Length 354;

Best Local Similarity 92.1%; Pred. No. 3.3e-96;
Matches 303; Conservative 15; Mismatches 11; Indels 0; Gaps 0;

Qy 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 60
Db 2 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 61

Qy 61 IVEIYGPESSGKTTTLQVIAAAQREGKTCAPIDAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 62 IVEIYGPESSGKTTTLQVIAAAQREGKTCAPIDAEHALDPIYARKLGVDIDNLLCSQPD 121

Qy 121 TGEQALEICDALARGAVDVIIVDSVAALTTPKAEIEGEIGDSHMGGLAARMMSQAMRKLKAG 180
Db 122 TGEQALEICDALTRSGAVDVIIVDSVAALTTPKAEIEGEIGDSHMGGLAARMMSQAMRKLKAG 181

Qy 181 NLKQSNLTLLIFINQIRMKIGVFNPGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVVG 240
Db 182 NLKQSNLTLLIFINQIRMKIGVFNPGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVVG 241

Qy 241 ETRVVKVKNKTAAPFKQAEFOILYCEGINFYGVLDLGVKKEKLEKAGAWYSYKGEKIQ 300
Db 242 ETRVVKVKNKTAAPFKQAEFOILYCEGINFYGVLDLGVKKEKLEKAGAWYSYKGEKIQ 301

Qy 301 GKANATNWLKONPETAKEIEKKVRELLLS 329
Db 302 GKANSNVLYKENPKVAELDKKLRDMLLS 330

RESULT 7
S37586
recombination protein reca - Yersinia pestis
N;Alternate names: recombinase A
C;Species: Yersinia pestis
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C;Accession: S37586
submitted to the EMBL Data Library, October 1993
A;Description: Complete nucleotide sequence of Yersinia pestis reca gene.
A;Reference number: S37586
A;Accession: S37586
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-356 <R>
A;Cross-references: UNIPROT:P37858; UNIPARC:UPI0000170617; EMBL:X75336; NID:g406793; PID:37586
C;Genetics:
A;Gene: reca
C;Superfamily: recombination protein reca
C;Keywords: ATP; DNA binding; DNA recombination; DNA repair; nucleotide binding; P-loop;
F:67-74/Region: nucleotide-binding motif A (P-loop)
F:141-146/Region: nucleotide-binding motif B
F:7/3/Binding site: ATP (lys) #status predicted

Query Match 91.2%; Score 1530; DB 2; Length 356;
Best Local Similarity 91.9%; Pred. No. 4.6e-96;
Matches 305; Conservative 14; Mismatches 13; Indels 0; Gaps 0;

Qy 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 60
Db 2 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 61

Qy 61 IVEIYGPESSGKTTTLQVIAAAQREGKTCAPIDAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 62 IVEIYGPESSGKTTTLQVIAAAQREGKTCAPIDAEHALDPIYARKLGVDIDNLLCSQPD 121

Qy 121 TGEQALEICDALARGAVDVIIVDSVAALTTPKAEIEGEIGDSHMGGLAARMMSQAMRKLKAG 180
Db 122 TGEQALEICDALTRSGAVDVIIVDSVAALTTPKAEIEGEIGDSHMGGLAARMMSQAMRKLKAG 181

Qy 181 NLKQSNLTLLIFINQIRMKIGVFNPGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVVG 240
Db 182 NLKQSNLTLLIFINQIRMKIGVFNPGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVVG 241

Qy 241 ETRVVKVKNKTAAPFKQAEFOILYCEGINFYGVLDLGVKKEKLEKAGAWYSYKGEKIQ 300

Db 242 ETRVVKVKNKTAAPFKQAEFOILYCEGINFYGVLDLGVKKEKLEKAGAWYSYKGEKIQ 301

Qy 301 GKANATNWLKONPETAKEIEKKVRELLLSNPN 332
Db 302 GKANASNYLKENPANAAELDKKLRMLNGN 333

RESULT 8
R0EBPM
recombination protein reca - Proteus mirabilis
N;Alternate names: recombinase A
C;Species: Proteus mirabilis
C;Date: 30-Sep-1990 #sequence_revision 30-Sep-1990 #text_change 09-Jul-2004
C;Accession: S04606
R;Akaboshi, E.; Yip, M.L.R.; Howard-Flanders, P.
Nucleic Acids Res. 17, 4390, 1989
A;Title: Nucleotide sequence of the reca gene of Proteus mirabilis.
A;Reference number: S04606; MUID:89296502; PMID:2544862
A;Accession: S04606
A;Molecule type: DNA
A;Residues: 1-355 <AKA>
A;Cross-references: UNIPROT:P11406; UNIPARC:UPI000016FD85; GB:X14870; NID:g45631; PIDN:355
C;Genetics:
A;Gene: reca
C;Function:
A;Description: plays an essential role in homologous recombination, in induction of the
C;Superfamily: recombination protein reca
C;Keywords: ATP; DNA binding; DNA recombination; DNA repair; nucleotide binding; P-loop;
F:67-74/Region: nucleotide-binding motif A (P-loop)
F:141-146/Region: nucleotide-binding motif B
F:7/3/Binding site: ATP (lys) #status predicted

Query Match 89.6%; Score 1504; DB 1; Length 355;
Best Local Similarity 89.4%; Pred. No. 2.6e-94;
Matches 295; Conservative 21; Mismatches 14; Indels 0; Gaps 0;

Qy 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 60
Db 2 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 61

Qy 61 IVEIYGPESSGKTTTLQVIAAAQREGKTCAPIDAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 62 IVEIYGPESSGKTTTLQVIAAAQREGKTCAPIDAEHALDPIYARKLGVDIDNLLCSQPD 121

Qy 121 TGEQALEICDALARGAVDVIIVDSVAALTTPKAEIEGEIGDSHMGGLAARMMSQAMRKLKAG 180
Db 122 TGEQALEICDALTRSGAVDVIIVDSVAALTTPKAEIEGEIGDSHMGGLAARMMSQAMRKLKAG 181

Qy 181 NLKQSNLTLLIFINQIRMKIGVFNPGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVVG 240
Db 182 NLKQSNLTLLIFINQIRMKIGVFNPGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVVG 241

Qy 241 ETRVVKVKNKTAAPFKQAEFOILYCEGINFYGVLDLGVKKEKLEKAGAWYSYKGEKIQ 300
Db 242 ETRVVKVKNKTAAPFKQAEFOILYCEGINFYGVLDLGVKKEKLEKAGAWYSYKGEKIQ 301

Qy 301 GKANATNWLKONPETAKEIEKKVRELLLSN 330
Db 302 GKANATNLYKHPPEWYNELNKLREMLNH 331

RESULT 9
B82310
reca protein VC0543 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C;Accession: B82310
R;Heideberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, F.
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: A82035, MUID:20406833, PMID:10952301

A:Accession: B82310

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-412 <HEI>

A:Cross-references: UNIPARC:UPI0000164B55; GB:AE004140; GB:AE003852; NID:g9654965; PIDN:

A:Experimental source: serogroup O1; strain N16961; biotype El Tor

C:Genetics:

A:Gene: VC0543

A:Map position: 1

C:Superfamily: recombination protein recA

Query Match 85.6%; Score 1436; DB 2; Length 412;

Best Local Similarity 84.3%; Pred. No. 1.2e-89;

Matches 279; Conservative 31; Mismatches 21; Indels 0; Gaps 0;

Qy 2 IDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMDVETISTGSLSLDIALGAGGLPMGRI 61

Db 59 MDENKQKALAAALGQIEKQFGKGSIMRLGDNRMVDVETISTGSLSLDIALGAGGLPMGRI 118

Qy 62 VEIYGPESSGKTTTLTQVIAAAQREGKTCAPIDAEHALDPYARKLVGDIDNLLCSQDPT 121

Db 119 VEIYGPESSGKTTTLTLELIAAAQREGKTCAPIDAEHALDPYAKLVGNIDELLVSQDPT 178

Qy 122 GEOALEICDALARSGAVDVIVDSVAALTPKAEITEGEIGDSHMGLAARMMSQAMRKLAGN 181

Db 179 GEOALEICDALARSGAVDVIVDSVAALTPKAEITEGEIGDSHMGLOARMLSQAMRKLTGN 238

Qy 182 LKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGSE 241

Db 239 LKQSNCMCIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAIKEGEVVGNE 298

Qy 242 TRVKVKNKIAAPPKQAEFOILYEGEINFYGELVDLVGKVKLIEKAGAWYSYKGEKIQG 301

Db 299 TRIKVKNKIAAPPKAEANTQIMYGQGFNREGELDLGVKHKWVEKSGAWSYNGDKIQG 358

Qy 302 KANATAMLKONPETAKEIEKKVRELLLSNP 332

Db 359 KANAYKVLKENPEIAKTLDKKREMLLNPN 389

RESULT 10

S46274

recombination protein recA - Vibrio cholerae

N:Alternate names: recombinase A

C:Species: Vibrio cholerae

C:Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 02-Feb-2001

C:Accession: S46274

R:Stroeder, U.H.; Lech, A.J.; Manning, P.A.

Mol. Gen. Genet. 244, 295-302, 1994

A:Title: Gene sequence of recA(+) and construction of recA mutants of Vibrio cholerae.

A:Reference number: S46274; MUID:94335880; PMID:8058040

A:Accession: S46274

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-354 <STR>

A:Cross-references: UNIPARC:UPI0000170586; EMBL:X71969; NID:g530269; PIDN:CAA50764.1; PI

C:Genetics:

A:Gene: recA

C:Superfamily: recombination protein recA

C:Keywords: ATP; DNA binding; DNA recombination; DNA repair; nucleotide binding; P-loop;

F:65-72/Region: nucleotide-binding motif A (P-loop)

F:139-144/Region: nucleotide-binding motif B

F:71/Binding site: ATP (lys) #status predicted

Query Match 84.4%; Score 1417; DB 2; Length 354;

Best Local Similarity 83.1%; Pred. No. 1.9e-88;

Matches 275; Conservative 31; Mismatches 25; Indels 0; Gaps 0;

Qy 2 IDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMDVETISTGSLSLDIALGAGGLPMGRI 61

Db 1 MDENKQKALAAALGQIEKQFGKGSIMHLDGNRMVDVETISTGSLSLDITLGGSLPMGRI 60

Qy 62 VEIYGPESSGKTTTLTQVIAAAQREGKTCAPIDAEHALDPYARKLVGDIDNLLCSQDPT 121

Db 61 VEIYGPESSGKTTTLTLELIAAAQREGKTCAPIDTEHALDPYAKLVGNIDELLVSQDPT 120

Qy 122 GEOALEICDALARSGAVDVIVDSVAALTPKAEITEGEIGDSHMGLAARMMSQAMRKLAGN 181

Db 121 GEOALEICDALARSGAVDVIVDSVAALTPKAEITEGEIGDSHMGLOARMLSQAMRKLTGN 180

Qy 182 LKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGSE 241

Db 181 LKQSNCMCIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAIKEGEVVGNE 240

Qy 242 TRVKVKNKIAAPPKQAEFOILYEGEINFYGELVDLVGKVKLIEKAGAWYSYKGEKIQG 301

Db 241 TRIKVKVKNKIAAPPKAEANTQIMYGQGFNREGELDLGVKHKWVEKSGAWSYNGDKIQG 300

Qy 302 KANATAMLKONPETAKEIEKKVRELLLSNP 332

Db 301 KANACKYVLKENPEIAKTLDKKREMLLNPN 331

RESULT 11

QJ1461

recombination protein recA - Vibrio anguillarum

N:Alternate names: recombinase A

C:Species: Vibrio anguillarum

C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 09-Jul-2004

C:Accession: QJ1461

R:Tolmashy, M.E.; Gammie, A.E.; Crosa, J.H.

Gene 110, 41-48, 1992

A:Title: Characterization of the recA gene of Vibrio anguillarum.

A:Reference number: QJ1461; MUID:92184113; PMID:1544576

A:Molecule type: DNA

A:Residues: 1-348 <TOL>

A:Cross-references: UNIPROT:P26348; UNIPARC:UPI0000133544; GB:M80525; NID:g155256; PIDN:

A:Experimental source: strains 775 and 531A

C:Genetics:

A:Gene: recA

C:Function:

C:Description: plays an essential role in homologous recombination, in induction of the

C:Superfamily: recombination protein recA

C:Keywords: ATP; DNA binding; DNA recombination; DNA repair; nucleotide binding; P-loop

F:65-72/Region: nucleotide-binding motif A (P-loop)

F:139-144/Region: nucleotide-binding motif B

F:256-279/Region: ATP binding #status predicted

F:71/Binding site: ATP (lys) #status predicted

Query Match 84.2%; Score 1412.5; DB 2; Length 348;

Best Local Similarity 83.6%; Pred. No. 3.8e-88;

Matches 276; Conservative 30; Mismatches 23; Indels 1; Gaps 1;

Qy 2 IDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMDVETISTGSLSLDIALGAGGLPMGRI 61

Db 1 MDENKQKALAAALGQIEKQFGKGSIMRLGDNRMVDVETISTGSLSLDIALGAGGLPMGRI 60

Qy 62 VEIYGPESSGKTTTLTQVIAAAQREGKTCAPIDAEHALDPYARKLVGDIDNLLCSQDPT 121

Db 61 VEIYGPESSGKTTTLTLELIAAAQREGKTCAPIDAEHALDPYAKLVGNIDELLVSQDPT 120

Qy 122 GEOALEICDALARSGAVDVIVDSVAALTPKAEITEGEIGDSHMGLAARMMSQAMRKLAGN 181

Db 121 GEOALEICDALARSGAIDVIDVSAALTPKAEITEGEIGDSHMGLOARMLSQAMRKLTGN 180

Qy 182 LKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGSE 241

Db 181 LKQSNCMCIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAIKEGEVVGNE 240

Qy 242 TRVKVKNKIAAPPKQAEFOILYEGEINFYGELVDLVGKVKLIEKAGAWYSYKGEKIQG 301

Db 241 TRIKVKVKNKIAAPPKQADTQILYGGQFNREGELVDLVGKHKLVEKAGAWSYNGDKIQG 300

Qy 302 KANATAMLKONPETAKEIEKKVRELLLSNP 331

```
||||| :|:| | :|:| | | :|:| | | |
Db 301 KANACKFLRENPAARAWALDKLREMLL-NP 329

RESULT 12
A49929
recombination protein recA - Haemophilus influenzae (strain Rd KW20)
N:Alternate names: rec-1 protein; recombinase A
C:Species: Haemophilus influenzae
C>Date: 11-Nov-1994 #sequence_revision 11-Nov-1994 #text_change 09-Jul-2004
C:Accession: A49929; A64080
R:Zulty, J.J.; Barcak, G.J.
J. Bacteriol. 175, 7269-7281, 1993
A:Title: Structural organization, nucleotide sequence, and regulation of the Haemophilus
A:Reference number: A49929; MUID:94042901; PMID:8228674
A:Accession: A49929
A:Molecule type: DNA
A:Residues: 1-354 <ZUL>
A:Cross-references: UNIPROT:P43705; UNIPARC:UPI00001334E8; GB:L07529; NID:G305381; PIDN:
A:Experimental source: Rd strain KW20
R:Flerschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J
, D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A:Reference number: A64000; MUID:95350630; PMID:7542800
A:Accession: A64080
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-354 <TIGR>
A:Cross-references: UNIPARC:UPI00001334E8; GB:U32741; GB:L42023; NID:g1573582; PIDN:AAC2
A:Experimental source: Rd strain KW20
C:Genetics:
A:Gene: recA
C:Function:
A:Description: plays an essential role in homologous recombination, in induction of the
C:Superfamily: recombination protein recA
C:Keywords: ATP; DNA binding; DNA recombination; DNA repair; nucleotide binding; P-loop;
F:67-74/Region: nucleotide-binding motif A (P-loop)
F:141-146/Region: nucleotide-binding motif B
F:73/Binding site: ATP (Lys) #status predicted

Query Match 79.4%; Score 1332; DB 2; Length 354;
Best Local Similarity 76.4%; Pred. No. 1.1e-82;
Matches 253; Conservative 42; Mismatches 36; Indels 0; Gaps 0;

Qy 1 ATDENKQKALAAALGOIEKQFGKSGIMRLGEDRSMDVETISTGSLSLDIALGAGGLPMGR 60
Db 2 ATQEQKQKALAAALGOIEKQFGKSGIMKLGDTKTLDVRSISTGSLGLDVALGIGLPMGR 61
Qy 61 IVEIYGPSSSGKTTTLTVIAAAQREGKTCAPIDAEHALDPIYARKLGVVDINLLCSQPD 120
Db 62 IVEIYGPSSSGKTTTLTVIAQAQKAGTKCAPIDAEHALDPIYAAKLGVDVVKELFVSQPD 121
Qy 121 TQEQALEICDALARSGADVIVVDSVAALTPKAEITEGIDSHMGLAARMWSQAMRKLAG 180
Db 122 NGEQALEICDALARSGADVIVVDSVAALTPKAEITEGDMGDSHMGLOARLMSQALRKUTG 181
Qy 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
Db 182 QIKNANCLVFINQIRMKIGVMFGNPETTTGGNALKFYSSVRLDIRRTGSGVVDGNIIGN 241
Qy 241 ETRVKVKNKIAAPKQAEFQILYEGEINFYGELVDLVGKELIEKAGAWYSYKGEKIQG 300
Db 242 ETRVKVKNKLAAPPRQVDFQILYEGEISKAGELLELVGKHLVEKSGAWTSYNGEKIQG 301
Qy 301 GKANATAMLKONPETAKEIEKKVRELLISNP 331
Db 302 GKANSKMKWLNENIEKSDLEALRLAELVANP 332

RESULT 13
```

```
JC5198
recombination protein recA - Aeromonas salmonicida
N:Alternate names: recombinase A
C:Species: Aeromonas salmonicida
C>Date: 27-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 09-Jul-2004
C:Accession: JC5198
R:Umelo, E.; Noonan, B.; Trust, T.J.
Gene 175, 133-136, 1996
A:Title: Cloning, characterization and expression of the recA gene of Aeromonas salmonic
A:Reference number: JC5198; MUID:97074662; PMID:8917089
A:Accession: JC5198
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-353 <UME>
A:Cross-references: UNIPROT:P94190; UNIPARC:UPI00001334C5; GB:U83688; NID:g1785949; PIDN:
C:Genetics:
A:Gene: recA
C:Function:
A:Description: plays an essential role in homologous recombination, in induction of the
C:Superfamily: recombination protein recA
C:Keywords: ATP; DNA binding; DNA recombination; DNA repair; nucleotide binding; P-loop;
F:65-72/Region: nucleotide-binding motif A (P-loop)
F:139-144/Region: nucleotide-binding motif B
F:71/Binding site: ATP (Lys) #status predicted

Query Match 78.8%; Score 1323; DB 2; Length 353;
Best Local Similarity 79.3%; Pred. No. 4.3e-82;
Matches 260; Conservative 30; Mismatches 38; Indels 0; Gaps 0;

Qy 2 IDENKQKALAAALGOIEKQFGKSGIMRLGEDRSMDVETISTGSLSLDIALGAGGLPMGRI 61
Db 1 MDQNKQKALAAALGOIEKQFGKSGIMLLGDSKTDWIDEAISTGSLSLDVALGIGLPCGRI 60
Qy 62 VEIYGPSSSGKTTTLTVIAAAQREGKTCAPIDAEHALDPIYARKLGVVDINLLCSQPD 121
Db 61 VEIYGPSSSGKTTTLTVIAEAQKGVKCAPIDAEHALDPIYAAKLGYNVVDLLISQSDT 120
Qy 122 GEQALEICDALARSGADVIVVDSVAALTPKAEITEGIDSHMGLAARMWSQAMRKLAGN 181
Db 121 GEQALEICDMLVRNSADVIVVDSVAALTPKAEIEGEGMDSHVGLQARLMSQALRKLTAN 180
Qy 182 LKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGSE 241
Db 181 IKVANCLCIPINQIRMKIGVMFGSPETTTGGNALKFYASVRLDIRRIGAIKEGDEVWGNE 240
Qy 242 TRVKVKNKIAAPKQAEFQILYEGEINFYGELVDLVGKELIEKAGAWYSYKGEKIQG 301
Db 241 TRVKVKNKIVAPPKQAEFQIFYGVGISKEGELVDLVGKHKLIDKAGAWYSYNGEKIQG 300
Qy 302 KANATAMLKONPETAKEIEKKVRELLIS 329
Db 301 KANVKKLFTENKVNAAVEEARLRELLIS 328

RESULT 14
RQPSAA
RecA protein PA3617 [imported] - Pseudomonas aeruginosa (strain PAO1)
N:Alternate names: recombinase A
C:Species: Pseudomonas aeruginosa
C>Date: 30-Sep-1990 #sequence_revision 30-Sep-1990 #text_change 09-Jul-2004
C:Accession: S06265; S10458; A49854; JN0304; G83192
R:Sano, Y.; Kageyama, M.
Mol. Gen. Genet. 208, 412-419, 1987
A:Title: The sequence and function of the recA gene and its protein in Pseudomonas aeru
A:Reference number: S06265; MUID:88038334; PMID:2823059
A:Accession: S06265
A:Molecule type: DNA
A:Residues: 1-346 <SAN1>
A:Cross-references: UNIPROT:P08280; UNIPARC:UPI0000133516; GB:X05691; NID:g45413; PIDN:
A>Note: part of this sequence, including the amino end of the mature protein, was confir
R:Zaitsev, E.N.; Krjukov, V.M.; Kuzmin, N.P.; Alekseev, A.A.; Lanzov, V.A.
submitted to the EMBL Data Library, March 1990
A:Description: Restriction polymorphism and nucleotide sequence substitutions in the rec
```


A:Reference number: S10458
A:Accession: S10458
A:Molecule type: DNA
A:Residues: 1-346 <ZAI>
A:CROSS-references: UNIPARC:UPI0000133516; EMBL:X52261; NID:G45383; PIDN:CAA36504.1; PID: R.Sano, Y.
J. Bacteriol. 175, 2451-2454, 1993
A:Title: Role of the recA-related gene adjacent to the recA gene in *Pseudomonas aeruginosa*
A:Reference number: A49854; MUID:93224470; PMID:8468303
A:Accession: A49854
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 313-346 <SAN2>
A:CROSS-references: UNIPARC:UPI00001748B1; GB:D13090; NID:G286185
A:Note: sequence extracted from NCBI backbone (NCBI:129317, NCBI:P:129318)
R.Kryukov, V.M.; Zaitsev, E.N.; Kouzmin, N.P.; Bayev, A.A.
Bioorg. Khim. 16, 1177-1182, 1990
A:Title: Structure of the recA gene from *Pseudomonas aeruginosa*.
A:Reference number: JN0304; MUID:91182156; PMID:2127886
A:Accession: JN0304
A:Molecule type: DNA
A:Residues: 1-332, 'L', 334-346 <KRY>
A:CROSS-references: UNIPARC:UPI00001748B2; GB:X52261; NID:G45383; PIDN:CAA36504.1; PID: R.Sano, Y.
A:Experimental source: strain PAW 7
A:Note: the authors translated the codon CTG for residue 333 as Val
R.Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B. adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, .; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen
A:Reference number: AB2950; MUID:20437337; PMID:10984043
A:Accession: G83192
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-346 <STO>
A:CROSS-references: UNIPARC:UPI0000133516; GB:AE004782; GB:AE004091; NID:G9949772; PIDN: R.Sano, Y.
A:Experimental source: strain PA01
C:Genetics:
A:Gene: recA; PA3617
C:Function:
A:Description: plays an essential role in homologous recombination, in induction of the SOS response
C:Superfamily: recombination protein recA
C:Keywords: ATP; DNA binding; DNA recombination; DNA repair; nucleotide binding; P-loop; F:65-72/Region: nucleotide-binding motif A (P-loop)
F:139-144/Region: nucleotide-binding motif B
F:71/Binding site: ATP (lys) #status predicted
Query Match 75.6%; Score 1268; DB 1; Length 346;
Best Local Similarity 74.4%; Pred. No. 2.1e-78;
Matches 244; Conservative 42; Mismatches 42; Indels 0; Gaps 0;
Qy 2 IDENKQKALAAALGOIEKQKFGKSGIMRLGEDRSDMDVETISTGSLDIALGAGGLPMGRI 61
Db 1 MDENKKALAAALGOIEKQKFGKSGIMRLGEDRSDMDVETISTGSLDIALGAGGLPMGRI 60
Qy 62 VEIYGPSSGKTTTLTVIAAAQREGKTCAPIDAEHALDPTIYARKLGVDIDNLLCSQPD 121
Db 61 VEIYGPSSGKTTTLTVIAAAQREGKTCAPIDAEHALDPTIYARKLGVDIDNLLCSQPD 120
Qy 122 GEOALEICDALARSGAVDVIIVDSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKLAGN 181
Db 121 GEOALEITDMLVRNAVDVIIVDSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKLAGN 180
Qy 182 LKQSNLTLLIFNQIRMKIGVMFGNPETTTGGNALKKFYASVRLDIRRGAVKESDEVV 241
Db 181 IKNANCLVIFNQIRMKIGVMFGNPETTTGGNALKKFYASVRLDIRRGAVKESDEVV 240
Qy 242 TRVKVKNKIAAPKQAEFQILYEGEINFYGLVGLVKEKLEKAGAWSYKGEKIGQG 301
Db 241 TRVKVKNKVPFQAEFQILYEGEINFYGLVGLVKEKLEKAGAWSYKGEKIGQG 300
Qy 302 KANATAMLKONPETAKIEKKVRELLS 329
Db 302 KANATAMLKONPETAKIEKKVRELLS 329

Db 301 KANAAKYLEDNPEIGSVLEKTRDQLLA 328
RESULT 15
JN0321
recombination protein recA - *Azotobacter vinelandii*
N:Alternate names: recombinase A
C:Species: *Azotobacter vinelandii*
C:Date: 09-Oct-1992 #sequence_revision 09-Oct-1992 #text_change 09-Jul-2004
C:Accession: JN0321
R:Kenkatesh, T.V.; Das, H.K.
Gene 113, 47-53, 1992
A:Title: The *Azotobacter vinelandii* recA gene: sequence analysis and regulation of expression
A:Reference number: JN0321; MUID:92225347; PMID:1563632
A:Accession: JN0321
A:Molecule type: DNA
A:Residues: 1-349 <KEN>
A:CROSS-references: UNIPROT:P29246; UNIPARC:UPI00001334CC
A:Note: the authors translated the codon GAC for residue 129 as Val and AAC for residue 130
C:Genetics:
A:Gene: recA
C:Function:
A:Description: plays an essential role in homologous recombination, in induction of the SOS response
C:Superfamily: recombination protein recA
C:Keywords: ATP; DNA binding; DNA recombination; DNA repair; DNA replication; nucleotide binding; P-loop; F:65-72/Region: nucleotide-binding motif A (P-loop)
F:139-144/Region: nucleotide-binding motif B
F:71/Binding site: ATP (lys) #status predicted
Query Match 75.4%; Score 1265; DB 2; Length 349;
Best Local Similarity 73.5%; Pred. No. 3.5e-78;
Matches 244; Conservative 41; Mismatches 47; Indels 0; Gaps 0;
Qy 2 IDENKQKALAAALGOIEKQKFGKSGIMRLGEDRSDMDVETISTGSLDIALGAGGLPMGRI 61
Db 1 MDENKKALAAALGOIEKQKFGKSGIMRLGEDRSDMDVETISTGSLDIALGAGGLPMGRI 60
Qy 62 VEIYGPSSGKTTTLTVIAAAQREGKTCAPIDAEHALDPTIYARKLGVDIDNLLCSQPD 121
Db 61 VEIYGPSSGKTTTLTVIAAAQREGKTCAPIDAEHALDPTIYARKLGVDIDNLLCSQPD 120
Qy 122 GEOALEICDALARSGAVDVIIVDSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKLAGN 181
Db 121 GEOALEITDMLVRNAVDVIIVDSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKLAGN 180
Qy 182 LKQSNLTLLIFNQIRMKIGVMFGNPETTTGGNALKKFYASVRLDIRRGAVKESDEVV 241
Db 181 IKNANCLVIFNQIRMKIGVMFGNPETTTGGNALKKFYASVRLDIRRGAVKESDEVV 240
Qy 242 TRVKVKNKIAAPKQAEFQILYEGEINFYGLVGLVKEKLEKAGAWSYKGEKIGQG 301
Db 241 TRVKVKNKVPFQAEFQILYEGEINFYGLVGLVKEKLEKAGAWSYKGEKIGQG 300
Qy 302 KANATAMLKONPETAKIEKKVRELLS 333
Db 301 KANAAKPLEDNPEVAAAVEKSIKRDQLLAAPAS 332
Search completed: February 16, 2006, 01:24:56
Job time : 26.5 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: February 16, 2006, 01:14:35 ; Search time 152 Seconds
(without alignments)
1554.947 Million cell updates/sec

Title: US-10-733-782-1

Perfect score: 1678

Sequence: 1 AIDENKQKALAAALGQIEKQ.....AKEIEKKVRELLSNPNSTP 335

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot_05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1678	100.0	352	1 RECA_ECO57	P0A798 escherichia
2	1678	100.0	352	1 RECA_ECOL6	P0A797 escherichia
3	1678	100.0	352	1 RECA_ECOL1	P0A796 escherichia
4	1678	100.0	352	1 RECA_SHIFL	P0A799 shigella fl
5	1671	99.6	352	1 RECA_SHISO	Q95ff6 shigella so
6	1641	97.8	352	1 RECA_SALTY	P65978 salmonella
7	1641	97.8	352	1 RECA_SALTY	P65977 salmonella
8	1641	97.8	352	2 Q57KU4_SALCH	Q57ku4 salmonella
9	1639	97.7	353	2 Q5PF15_SALPA	Q5pf15 salmonella
10	1547	92.2	355	1 RECA_PHOLL	Q7n7a5 photorhabdu
11	1537	91.6	357	1 RECA_ERWCT	Q6dl88 erwinia car
12	1534	91.4	355	1 RECA_YERPE	P37858 yersinia pe
13	1534	91.4	356	1 RECA_YERPS	Q66e70 yersinia ps
14	1532	91.3	353	1 RECA_ENTAG	P33037 enterobacta
15	1527	91.0	353	1 RECA_SERMA	P17479 serratia ma
16	1524	90.8	358	1 RECA_XENRV	P96185 xenorhabdus
17	1523	90.8	355	1 RECA_SODGL	P62220 sodalis glo
18	1513	90.2	355	2 Q5Y131_9ENTR	Q5y131 primary end
19	1504	89.6	354	1 RECA_PROMI	P11406 proteus mir
20	1501	89.5	342	1 RECA_ERWCA	P26344 erwinia car
21	1501	89.5	357	1 RECA_SHEON	Q8eb50 shewanella
22	1500	89.4	358	1 RECA_XENNE	Q9X5p5 xenorhabdus
23	1486	88.6	325	1 RECA_PROVU	P26346 proteus vul
24	1437	85.6	354	1 RECA_VIBCH	P45383 vibrio chol
25	1429	85.2	347	1 RECA_VIBPA	Q871r1 vibrio para
26	1418	84.5	348	1 RECA_VIBNA	Q6xz07 vibrio natr
27	1418	84.5	349	1 RECA_VIBVU	Q8dc51 vibrio vuln
28	1418	84.5	349	1 RECA_VIBVY	Q7mhr4 vibrio vuln
29	1416	84.4	348	2 Q5E7G6_VIBFI	Q5e7g6 vibrio fisc
30	1412.5	84.2	348	1 RECA_VIBAN	P26348 vibrio angu
31	1404	83.7	352	1 RECA_PHOPR	Q61mu2 photobacter

32 1400 83.4 346 2 Q9S4R6_VIBCH Q9s4r6 vibrio chol
33 1400 83.4 346 2 Q9R2W9_VIBCH Q9r2w9 vibrio chol
34 1353 80.6 354 1 RECA_PASWU P95526 pasteurella
35 1332 79.4 354 1 RECA_HAEIN P43705 haemophilus
36 1332 79.4 354 2 Q4QW2_HAEI8 Q4qmw2 haemophilus
37 1326 79.0 348 1 RECA_PSEOL Q9r9u2 pseudomonas
38 1326 79.0 372 2 Q5QUB8_IDILO Q5qub8 idiomarina
39 1323 78.8 353 1 RECA_AERSA P94190 aeromonas s
40 1319 78.6 351 1 RECA_MANSM Q65qb0 mannhelma
41 1317 78.5 331 1 RECA_WIGBR Q8d2w7 wigglewort
42 1316 78.4 352 1 RECA_ACTAC Q9jrp9 actinobacill
43 1288 76.8 368 1 RECA_PASHA Q9rny0 pasteurella
44 1276 76.0 347 2 Q6EV36_PSEST Q6ev36 pseudomonas
45 1275 76.0 349 2 Q4IXX9_AZOVI Q4ixx9 azotobacter

ALIGNMENTS

RESULT 1

RECA_ECO57 STANDARD; PRT; 352 AA.
AC P0A7G8; P03017; P26347; P78213;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE RECA protein (Recombinase A).
GN Name=RECA; OrderedLocusNames=z4002, EC83556;
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=O157:H7 / EDJ933 / ATCC 700927 / EHEC;
RX MEDLINE=21074935; PubMed=11206551; DOI=10.1038/35054089;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grotbeck R.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";
RL Nature 409:529-533(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=O157:H7 / Sakai / RIMD 0509952 / EHEC;
RX MEDLINE=21156231; PubMed=11256796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
CC -!- FUNCTION: Can catalyze the hydrolysis of ATP in the presence of
single-stranded DNA, the ATP-dependent uptake of single-stranded
DNA by duplex DNA, and the ATP-dependent hybridization of
homologous single-stranded DNAs. It interacts with IexA causing
its activation and leading to its autocatalytic cleavage (By
similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the RECA family.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
CC EMBL; AE005174; AG57804.1; -; Genomic DNA.
DR EMBL; BA000007; BAB36979.1; -; Genomic DNA.
DR PIR; D91073; D91073.

DR HAMAP; MF_00268; -; 1.
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR001553; RecA.
 DR Pfam; PF00154; RecA; 1.
 DR PRINTS; PR00142; RecA.
 DR ProDom; PD000229; RecA; 1.
 DR SMART; SM00382; AAA; 1.
 DR TIGRFAMs; TIGR02012; tigrfam_recA; 1.
 DR PROSITE; PS00321; RecA_1; 1.
 DR PROSITE; PS50162; RecA_2; 1.
 DR PROSITE; PS50163; RecA_3; 1.
 KW ATP-binding; Complete proteome; DNA damage; DNA recombination;
 KW ATP repair; DNA-binding; Nucleotide-binding; SOS response.
 FT INIT_MET 0 By similarity.
 FT NP_BIND 66 73 ATP (By similarity).
 FT CONFLICT 139 139 D -> N (in Ref. 1).
 SQ SEQUENCE 352 AA; 37842 MW; 989B02378EC4A402 CRC64;

Query Match 100.0%; Score 1678; DB 1; Length 352;
 Best Local Similarity 100.0%; Pred. No. 5.4e-104;
 Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AIDENKQKALAAALGOIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 60
 DB 1 AIDENKQKALAAALGOIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 60
 QY 61 IVEIYGPESSGKTTTLQVIAAAAREGKTCFADAEHALDPIYARKLGVDDNLLCSQPD 120
 DL 61 IVEIYGPESSGKTTTLQVIAAAAREGKTCFADAEHALDPIYARKLGVDDNLLCSQPD 120
 QY 121 TGEQALEICDALARGAVDVIVDSVAALTPKAEIEGIGSHMGLAARMMSQAMRKLAG 180
 DB 121 TGEQALEICDALARGAVDVIVDSVAALTPKAEIEGIGSHMGLAARMMSQAMRKLAG 180
 QY 181 NLKQNTLLIFINQIRMKIGVFNPGPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
 DB 181 NLKQNTLLIFINQIRMKIGVFNPGPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
 QY 241 ETRVKVKNKIAAPFKQAEFQILYEGEINFGYELVDLGVKEKLEKAGAWYSYKGEKIGQ 300
 DB 241 ETRVKVKNKIAAPFKQAEFQILYEGEINFGYELVDLGVKEKLEKAGAWYSYKGEKIGQ 300
 QY 301 GKANATAWLKONPETAKIEKKVRELLSNPNSTP 335
 DB 301 GKANATAWLKONPETAKIEKKVRELLSNPNSTP 335

RESULT 2
 RECA_ECOLI6 STANDARD; PRT; 352 AA.
 AC POA7G6; P03017; P26347; P78213;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE RecA protein (Recombinase A).
 GN Name=recA; OrderedLocusNames=c3253;
 OS Escherichia coli O6.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OX NCBI_TaxID=217992;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=O6:H1 / CFT073 / ATCC 700928 / UPEC;
 RX MEDLINE=22388234; PubMed=12471157; DOI=10.1073/pnas.252529799;
 RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
 RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
 RA Mayhew G.P., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
 RA Mobley H.L.F., Donnenberg M.S., Blattner F.R.;
 RT of uropathogenic Escherichia coli.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
 CC -!- FUNCTION: Can catalyze the hydrolysis of ATP in the presence of
 single-stranded DNA, the ATP-dependent uptake of single-stranded

CC DNA by duplex DNA, and the ATP-dependent hybridization of
 CC homologous single-stranded DNAs. It interacts with LexA causing
 CC its activation and leading to its autocatalytic cleavage (By
 CC similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- SIMILARITY: Belongs to the RecA family.
 CC
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC EMBL; AE016765; AA081704.1; -; Genomic_DNA.
 DR HAMAP; MF_00268; -; 1.
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR001553; RecA.
 DR Pfam; PF00154; RecA; 1.
 DR PRINTS; PR00142; RecA.
 DR ProDom; PD000229; RecA; 1.
 DR SMART; SM00382; AAA; 1.
 DR TIGRFAMs; TIGR02012; tigrfam_recA; 1.
 DR PROSITE; PS00321; RecA_1; 1.
 DR PROSITE; PS50162; RecA_2; 1.
 DR PROSITE; PS50163; RecA_3; 1.
 KW ATP-binding; Complete proteome; DNA damage; DNA recombination;
 KW ATP repair; DNA-binding; Nucleotide-binding; SOS response.
 FT INIT_MET 0 By similarity.
 FT NP_BIND 66 73 ATP (By similarity).
 SQ SEQUENCE 352 AA; 37842 MW; 989B02378EC4A402 CRC64;

Query Match 100.0%; Score 1678; DB 1; Length 352;
 Best Local Similarity 100.0%; Pred. No. 5.4e-104;
 Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AIDENKQKALAAALGOIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 60
 DB 1 AIDENKQKALAAALGOIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 60
 QY 61 IVEIYGPESSGKTTTLQVIAAAAREGKTCFADAEHALDPIYARKLGVDDNLLCSQPD 120
 DB 61 IVEIYGPESSGKTTTLQVIAAAAREGKTCFADAEHALDPIYARKLGVDDNLLCSQPD 120
 QY 121 TGEQALEICDALARGAVDVIVDSVAALTPKAEIEGIGSHMGLAARMMSQAMRKLAG 180
 DB 121 TGEQALEICDALARGAVDVIVDSVAALTPKAEIEGIGSHMGLAARMMSQAMRKLAG 180
 QY 181 NLKQNTLLIFINQIRMKIGVFNPGPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
 DB 181 NLKQNTLLIFINQIRMKIGVFNPGPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
 QY 241 ETRVKVKNKIAAPFKQAEFQILYEGEINFGYELVDLGVKEKLEKAGAWYSYKGEKIGQ 300
 DB 241 ETRVKVKNKIAAPFKQAEFQILYEGEINFGYELVDLGVKEKLEKAGAWYSYKGEKIGQ 300
 QY 301 GKANATAWLKONPETAKIEKKVRELLSNPNSTP 335
 DB 301 GKANATAWLKONPETAKIEKKVRELLSNPNSTP 335

RESULT 3
 RECA_ECOLI STANDARD; PRT; 352 AA.
 AC POA7G6; P03017; P26347; P78213;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE RecA protein (Recombinase A).
 GN Name=recA; Synonym=lexB; recH, rnmB, tif, umuB, zab;
 OS OrderedLocusNames=b2699;
 ON Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OX Enterobacteriaceae; Escherichia.


```

FT CONFLICT 112 112 D -> E (in Ref. 5).
FT CONFLICT 190 190 Missing (in Ref. 7).
FT TURN 4 4
FT HELIX 5 21
FT HELIX 23 25
FT TURN 29 30
FT STRAND 39 40
FT HELIX 45 50
FT TURN 51 51
FT STRAND 56 57
FT TURN 58 59
FT STRAND 61 65
FT TURN 68 69
FT HELIX 72 85
FT TURN 86 87
FT STRAND 90 94
FT HELIX 101 106
FT TURN 107 108
FT HELIX 111 113
FT STRAND 115 117
FT HELIX 122 135
FT STRAND 140 144
FT HELIX 146 148
FT HELIX 152 155
FT TURN 166 186
FT STRAND 188 193
FT HELIX 213 218
FT STRAND 221 233
FT TURN 234 235
FT STRAND 236 249
FT STRAND 257 263
FT TURN 264 266
FT STRAND 267 268
FT HELIX 270 280
FT TURN 281 282
FT STRAND 285 287
FT TURN 288 289
FT STRAND 290 293
FT TURN 294 295
FT STRAND 296 300
FT HELIX 301 311
FT HELIX 313 327

Query Match 100.0%; Score 1678; DB 1; Length 352;
Best Local Similarity 100.0%; Pred. No. 5.4e-104;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMDVETISTGSLSLDIALGAGGLPMGR 60
DB 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMDVETISTGSLSLDIALGAGGLPMGR 60

QY 61 IVEIYGPSSGKTTTLQVIAAARQEGKTCFIDEHALDPIYARKLGVDIDNLLCSQPD 120
DB 61 IVEIYGPSSGKTTTLQVIAAARQEGKTCFIDEHALDPIYARKLGVDIDNLLCSQPD 120

QY 121 TGEQALEICDALARSGVDVIIVDSVAALTTPKAEIEGEIGDSHMGLAARMMSQAMKLAG 180
DB 121 TGEQALEICDALARSGVDVIIVDSVAALTTPKAEIEGEIGDSHMGLAARMMSQAMKLAG 180

QY 181 NLKQNTLLIFINQIRMKIGVFNPGPETTTGGNALKFVASVRLDIRRIGAVKEGNNVGS 240
DB 181 NLKQNTLLIFINQIRMKIGVFNPGPETTTGGNALKFVASVRLDIRRIGAVKEGNNVGS 240

QY 241 ETRVVKVKNKIAPPKQAEFQLYEGEINFYGEINFGYELVDLGKVKELIEKAGAWSYKGEKIQ 300
DB 241 ETRVVKVKNKIAPPKQAEFQLYEGEINFYGEINFGYELVDLGKVKELIEKAGAWSYKGEKIQ 300

QY 301 GKANATAMWKONPETAKETIEKKVRELLLSNPNSTP 335
DB 301 GKANATAMWKONPETAKETIEKKVRELLLSNPNSTP 335

RESULT 4
RECA_SHIFL STANDARD; PRT; 352 AA.
ID RECA_SHIFL STANDARD; PRT; 352 AA.
AC POA7G9; P03017; P26347; P78213;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE RecA protein (Recombinase A).
DN Name=recA; OrderedLocNames=Sf2722, S2913;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=BS12 / Serotype 2a;
RX MEDLINE=91109725; PubMed=2274037;
RA Zhao X.J., McEntee K.;
RT "DNA sequence analysis of the recA genes from Proteus vulgaris,
RT Erwinia carotovora, Shigella flexneri and Escherichia coli B/r.";
RL Mol. Gen. Genet. 222:369-376(1990).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=301 / Serotype 2a;
RX MEDLINE=22272406; PubMed=12384590; DOI=10.1093/nar/gkf566;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157.";
RL Nucleic Acids Res. 30:4432-4441(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=2457T / ATCC 700930 / Serotype 2a;
RX MEDLINE=22590274; PubMed=12704152;
DOI=10.1128/IAI.71.5.2775-2786.2003;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of Shigella
RT flexneri serotype 2a strain 2457T.";
RL Infect. Immun. 71:2775-2786(2003).
CC -!- FUNCTION: Can catalyze the hydrolysis of ATP in the presence of
CC single-stranded DNA, the ATP-dependent uptake of single-stranded
CC DNA by duplex DNA, and the ATP-dependent hybridization of
CC homologous single-stranded DNAs. It interacts with LexA causing
CC its activation and leading to its autocatalytic cleavage (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the recA family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; X55553; CAB56806.1; -; Genomic DNA.
CC EMBL; AE005674; AAN44244.1; ALT_INIT; Genomic_DNA.
CC EMBL; AE016987; AAP18040.1; -; Genomic_DNA.
CC HANAP; MF_00268; -; 1.
CC InterPro; IPR003593; AAA_ATPase.
CC InterPro; IPR001553; RecA.
CC Pfam; PF00154; RecA. 1.
CC PRINTS; PR00142; RecA.
CC ProDom; PD000229; RecA; 1.
CC SMART; SM00382; AAA; 1.
CC TIGRFAMs; TIGR02012; tigrfam_recA; 1.
CC PROSITE; PS00321; RECA_1; 1.
CC PROSITE; PS50162; RECA_2; 1.

```

```

DR PROSITE, PS50163; RECA_3; 1.
KW ATP-binding; Complete proteome; DNA damage; DNA recombination;
KW DNA repair; DNA-binding; Nucleotide-binding; SOS response.
FT INIT MET 0 By similarity.
FT NP BIND 66 73 ATP (By similarity).
SQ SEQUENCE 352 AA; 37842 MW; 989B02378EC4A402 CRC64;

Query Match
Best Local Similarity 100.0%; Score 1678; DB 1; Length 352;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AIDENKQKALAAALGOIEKQFGKGSIMRLGEDRSMDVETISTGSLSLDIALGAGGLPMGR 60
DB 1 AIDENKQKALAAALGOIEKQFGKGSIMRLGEDRSMDVETISTGSLSLDIALGAGGLPMGR 60
QY 61 IVEIYGPESSGKTTLTQVIAAAREGKTCFIDAEHALDPTIYARKLGVDIDNLLCSQPD 120
DB 61 IVEIYGPESSGKTTLTQVIAAAREGKTCFIDAEHALDPTIYARKLGVDIDNLLCSQPD 120
QY 121 TGEQALEICDALARSGADVIVVDSVAALTPKAEIEGIGDSHMGLAARMMSQAMRKLK 180
DB 121 TGEQALEICDALARSGADVIVVDSVAALTPKAEIEGIGDSHMGLAARMMSQAMRKLK 180
QY 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVGS 240
DB 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVGS 240
QY 241 ETRVKVVKNTIAAPFKQAEFOLLGEGINFGYELVDLGVEKLEKAGAWYSYKGEKIQ 300
DB 241 ETRVKVVKNTIAAPFKQAEFOLLGEGINFGYELVDLGVEKLEKAGAWYSYKGEKIQ 300
QY 301 GKANATAWLKONPETAKEIEKKVRELLSNPNSTP 335
DB 301 GKANATAWLKONPETAKEIEKKVRELLSNPNSTP 335

RESULT 5
RECA SHISO STANDARD; PRT; 352 AA.
AC Q92FF6;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE RecA protein (Recombinase A).
GN Name=recA;
OS Shigella sonnei.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=624;
[1]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RC STRAIN=KNH104S;
RA Park Y.C., Shin H.J., Kim Y.C.;
RT "Cloning and nucleotide sequence of the recA gene from Shigella sonnei
  KNH104S."
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Can catalyze the hydrolysis of ATP in the presence of
CC single-stranded DNA, the ATP-dependent uptake of single-stranded
CC DNA by duplex DNA, and the ATP-dependent hybridization of
CC homologous single-stranded DNAs. It interacts with LexA causing
CC its activation and leading to its autocatalytic cleavage.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the recA family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AF101227; AAC72856.1; -; Genomic_DNA.
CC HSSP; P03017; 2REB.
CC SMR; Q92FF6; 3-328.

```

```

DR HAMAP; MF_00268; -; 1.
DR InterPro; IPR003593; AAA ATPase.
DR Pfam; PF00154; RecA; 1.
DR PRINTS; PR00142; RECA.
DR PRODOM; PD000229; RecA; 1.
DR SMART; SM00382; AAA; 1.
DR TIGRFAMs; TIGR02012; tigrfam_recA; 1.
DR PROSITE; PS00321; RECA_1; 1.
DR PROSITE; PS0162; RECA_2; 1.
DR PROSITE; PS0163; RECA_3; 1.
KW ATP-binding; DNA damage; DNA recombination; DNA repair; DNA-binding;
KW Nucleotide-binding; SOS response.
FT INIT MET 0 By similarity.
FT NP BIND 66 73 ATP (By similarity).
SQ SEQUENCE 352 AA; 37799 MW; EEC50231893AA40C CRC64;

Query Match
Best Local Similarity 99.6%; Score 1671; DB 1; Length 352;
Matches 334; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AIDENKQKALAAALGOIEKQFGKGSIMRLGEDRSMDVETISTGSLSLDIALGAGGLPMGR 60
DB 1 AIDENKQKALAAALGOIEKQFGKGSIMRLGEDRSMDVETISTGSLSLDIALGAGGLPMGR 60
QY 61 IVEIYGPESSGKTTLTQVIAAAREGKTCFIDAEHALDPTIYARKLGVDIDNLLCSQPD 120
DB 61 IVEIYGPESSGKTTLTQVIAAAREGKTCFIDAEHALDPTIYARKLGVDIDNLLCSQPD 120
QY 121 TGEQALEICDALARSGADVIVVDSVAALTPKAEIEGIGDSHMGLAARMMSQAMRKLK 180
DB 121 TGEQALEICDALARSGADVIVVDSVAALTPKAEIEGIGDSHMGLAARMMSQAMRKLK 180
QY 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVGS 240
DB 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVGS 240
QY 241 ETRVKVVKNTIAAPFKQAEFOLLGEGINFGYELVDLGVEKLEKAGAWYSYKGEKIQ 300
DB 241 ETRVKVVKNTIAAPFKQAEFOLLGEGINFGYELVDLGVEKLEKAGAWYSYKGEKIQ 300
QY 301 GKANATAWLKONPETAKEIEKKVRELLSNPNSTP 335
DB 301 GKANATAWLKONPETAKEIEKKVRELLSNPNSTP 335

RESULT 6
RECA SALTI STANDARD; PRT; 352 AA.
AC P65978; O8XET0;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE RecA protein (Recombinase A).
GN Name=recA; OrderedLocusNames=STV2950, t2730;
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=601;
[1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC MEDLINE=21534947; PubMed=1677608; DOI=10.1038/35101607;
RX STRAIN=CT18;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C.M., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M.A., Rutherford K.M., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrrell B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
  enterica serovar Typhi CT18."

```



```

RL Nature 413:848-852(2001).
RN (2)
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Ty2 / ATCC 700931;
RX MEDLINE=22531367; PubMed=12644504;
DOI=10.1128/JB.185.7.2330-2337.2003;
RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA Burland V., Kodoyanni V., Schwartz D.C., Blattner F.R.;
RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
and C718.";
RL J. Bacteriol. 185:2330-2337(2003).
CC -!- FUNCTION: Can catalyze the hydrolysis of ATP in the presence of
CC single-stranded DNA, the ATP-dependent uptake of single-stranded
CC DNA by duplex DNA, and the ATP-dependent hybridization of
CC homologous single-stranded DNAs. It interacts with lexA causing
CC its activation and leading to its autocatalytic cleavage (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the recA family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; AL627276; CAD05935.1; -; Genomic_DNA.
DR EMBL; AE016843; AAO70291.1; -; Genomic_DNA.
DR HSSP; P03017; 2REB.
DR SMR; P65978; 3-328.
DR HAMAP; MF_00268; 1.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001553; RecA.
DR Pfam; PF00154; RecA; 1.
DR PRINTS; PR00142; RecA.
DR ProDom; PD000229; RecA; 1.
DR SMART; SM00382; AAA; 1.
DR TIGRFAMs; TIGR02012; tigrfam_recA; 1.
DR PROSITE; PS00321; RECA_1; 1.
DR PROSITE; PS0162; RECA_2; 1.
DR PROSITE; PS0163; RECA_3; 1.
DR ATP-binding; Complete proteome; DNA damage; DNA recombination;
KW DNA repair; DNA-binding; Nucleotide-binding; SOS response.
FT INIT MET 0
FT NP_BIND 66 73 ATP (By similarity).
SQ SEQUENCE 352 AA; 37813 MW; F24E51CB9A31B818 CRC64;

Query Match 97.8%; Score 1641; DB 1; Length 352;
Best Local Similarity 97.3%; Pred. No. 1.6e-101;
Matches 326; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 AIDENKQALAAALGQIEKQFGKSGIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 60
DB 1 AIDENKQALAAALGQIEKQFGKSGIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 60
QY 61 IVEIYGPSSGKTTLTQVIAAAQREGKTCAPIDAHALDPIYARKLGVDIDNLLCSQPD 120
DB 61 IVEIYGPSSGKTTLTQVIAAAQREGKTCAPIDAHALDPIYARKLGVDIDNLLCSQPD 120
QY 121 TGEQALEICDALARSGAVDIVVDSVAALTPKAEIEGETGDSHMGLAARMMSQAMRKLAG 180
DB 121 TGEQALEICDALARSGAVDIVVDSVAALTPKAEIEGETGDSHMGLAARMMSQAMRKLAG 180
QY 181 NLKQNTLLIFNQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVYGS 240
DB 181 NLKQNTLLIFNQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVYGS 240
QY 241 ETRVKVKNKIAAPFKQAFQILYEGEINFYGELVDLGVKEKLEKAGAWYSYKEKIQG 300
DB 241 ETRVKVKNKIAAPFKQAFQILYEGEINFYGELVDLGVKEKLEKAGAWYSYKEKIQG 300
QY 301 GKANATAWLKNDPETAKEIEKRVRELLLSNQNPSTP 335
DB 301 GKANATAWLKNDPETAKEIEKRVRELLLSNQNPSTP 335

```

Db 61 IVEIYGPSSGKTTTLQVIAAAQREGKTCAFIDAEHALDPYARKLGVLDNLLCSQPD 120
QY 121 TGEQALEICDALARSGAVDVIVDSVAALTPKAEIEGIGDSHMGGLAARMMSQAMRKLKAG 180
Db 121 TGEQALEICDALARSGAVDVIVDSVAALTPKAEIEGIGDSHMGGLAARMMSQAMRKLKAG 180
QY 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVVG 240
Db 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVVG 240
QY 241 ETRVKVKNKTAAPFKQAEFOILYCEGINFYGELVDLGVKELIEKAGAWSYKGEKIQ 300
Db 241 ETRVKVKNKTAAPFKQAEFOILYCEGINFYGELVDLGVKELIEKAGAWSYKGEKIQ 300
QY 301 GKANATWLKONPETAKEIEKKVRELLLSNPNSTP 335
Db 301 GKANATWLKENPATAKEIEKKVRELLLSNQATP 335

RESULT 8
Q57KU4 SALCH
ID Q57KU4 SALCH PRELIMINARY; PRT; 353 AA.
AC Q57KU4
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE DNA strand exchange and recombination protein with protease and
nuclease activity
GN Name=recA; OrderedLocusNames=SC2762;
OS Salmonella choleraesuis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=591;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=SC-B67;
RX PubMed=15781495; DOI=10.1093/nar/gki297;
RA Chiu C.-H., Tang P., Chu C., Hu S., Bao Q., Yu J., Chou Y.-Y.,
Wang H.-S., Lee Y.-S.
RT "The genome sequence of Salmonella enterica serovar Choleraesuis, a
highly invasive and resistant zoonotic pathogen."
RL Nucleic Acids Res. 33:1690-1698(2005).
REMBL; AE017220; AAX66668.1; -; Genomic_DNA.
DR SMR; Q57KU4; 4-329.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003684; F:damaged DNA binding; IEA.
DR GO; GO:0008094; F:DNA-dependent ATPase activity; IEA.
DR GO; GO:0017111; F:nucleoside-triphosphatase activity; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR GO; GO:0006281; P:DNA repair; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001553; RecA.
DR Pfam; PF00154; RecA.
DR PRINTS; PR00142; RECA.
DR SMART; SMO0382; AAA.
DR TIGRFAMs; TIGR02012; tigrfam_recA; 1.
DR PROSITE; PS00321; RECA_1; 1.
DR PROSITE; PS0162; RECA_2; 1.
DR PROSITE; PS0163; RECA_3; 1.
KW Complete proteome.
SQ SEQUENCE 353 AA; 37944 MW; ECBBAB042F16362C CRC64;

Query Match 97.8%; Score 1641; DB 2; Length 353;
Best Local Similarity 97.3%; Pred. No. 1.6e-101;
Matches 326; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
QY 1 AIDENKOKALAAALGQIEKQFGKGSIMRLGDRSDMDVETITSGLSLDIALGAGGLPMGR 60
Db 2 AIDENKOKALAAALGQIEKQFGKGSIMRLGDRSDMDVETITSGLSLDIALGAGGLPMGR 61
QY 61 IVEIYGPSSGKTTTLQVIAAAQREGKTCAFIDAEHALDPYARKLGVLDNLLCSQPD 120

Db 62 IVEIYGPSSGKTTTLQVIAAAQREGKTCAFIDAEHALDPYARKLGVLDNLLCSQPD 121
QY 121 TGEQALEICDALARSGAVDVIVDSVAALTPKAEIEGIGDSHMGGLAARMMSQAMRKLKAG 180
Db 122 TGEQALEICDALARSGAVDVIVDSVAALTPKAEIEGIGDSHMGGLAARMMSQAMRKLKAG 181
QY 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVVG 240
Db 182 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVVG 241
QY 241 ETRVKVKNKTAAPFKQAEFOILYCEGINFYGELVDLGVKELIEKAGAWSYKGEKIQ 300
Db 242 ETRVKVKNKTAAPFKQAEFOILYCEGINFYGELVDLGVKELIEKAGAWSYKGEKIQ 301
QY 301 GKANATWLKONPETAKEIEKKVRELLLSNPNSTP 335
Db 302 GKANATWLKENPATAKEIEKKVRELLLSNQATP 336

RESULT 9
Q5PF15 SALPA
ID Q5PF15 SALPA PRELIMINARY; PRT; 353 AA.
AC Q5PF15;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE RecA protein.
GN Name=recA; OrderedLocusNames=SPA2687;
OS Salmonella paratyphi-a.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=54388;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 9150 / SARB42;
RX PubMed=15531882; DOI=10.1038/ngl470;
RA McClelland M., Sanderson K.E., Clifton S.W., Latreille P.,
Porwollik S., Sabo A., Meyer R., Bieri T., Ozeraky P., McLellan M.,
Harkins C.R., Wang C., Nguyen C., Berghoff A., Elliott G.,
Kohlberg S., Strong C., Du F., Carter J., Kremizki C., Layman D.,
Leonard S., Sun H., Fulton L., Nash W., Miner T., Minx P.,
Delehaunty K., Fronick C., Magrini V., Nhan M., Warren W., Florea L.,
Spieth J., Wilson R.K.
RT "Comparison of genome degradation in Paratyphi A and Typhi, human-
restricted serovars of Salmonella enterica that cause typhoid.";
RL Nat. Genet. 36:1268-1274(2004).
REMBL; CP000026; AAV78544.1; -; Genomic_DNA.
DR SMR; Q5PF15; 4-329.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003684; F:damaged DNA binding; IEA.
DR GO; GO:0008094; F:DNA-dependent ATPase activity; IEA.
DR GO; GO:0017111; F:nucleoside-triphosphatase activity; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR GO; GO:0006281; P:DNA repair; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001553; RecA.
DR Pfam; PF00154; RecA; 1.
DR PRINTS; PR00142; RECA.
DR ProDom; PD000229; RecA; 1.
DR SMART; SMO0382; AAA; 1.
DR TIGRFAMs; TIGR02012; tigrfam_recA; 1.
DR PROSITE; PS00321; RECA_1; 1.
DR PROSITE; PS0162; RECA_2; 1.
DR PROSITE; PS0163; RECA_3; 1.
KW Complete proteome.
SQ SEQUENCE 353 AA; 37962 MW; AD15A0C48E7CF6FD CRC64;

Query Match 97.7%; Score 1639; DB 2; Length 353;
Best Local Similarity 97.0%; Pred. No. 2.2e-101;
Matches 325; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

```
QY 1 AIDENKOKALAAALGOIEKQFGKSGIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 60
Db 2 AIDENKOKALAAALGOIEKQFGKSGIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 61
QY 61 IVEIYGPSSGKTTTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 62 IVEIYGPSSGKTTTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 121
QY 121 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGEGSHMGLAARMMSQAMRKLKAG 180
Db 122 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGEGSHMGLAARMMSQAMRKLKAG 181
QY 181 NLKQNTLLIFINQIRMKIGVMFGNPTTTGGNALKFYASVRLDIRRIGAVKEGNNVVG 240
Db 182 NLKQNTLLIFINQIRMKIGVMFGNPTTTGGNALKFYASVRLDIRRIGAVKEGNNVVG 241
QY 241 ETRVVKVKNKIAAPFKQAEFQILYEGEGINFYGGELVDLGKVKELIEKAGAWSYNGEKIGQ 300
Db 242 ETRVVKVKNKIAAPFKQAEFQILYEGEGINFYGGELVDLGKVKELIEKAGAWSYNGEKIGQ 301
QY 301 GKANATWLKONPETAKIEKKVRELLLSNPNSTP 335
Db 302 GKANATWLKONPETAKIEKKVRELLLSNQATP 336

RESULT 10
RECA_PHOLL
ID RECA PHOLL STANDARD; PRT; 355 AA.
AC Q7N7A6;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE RecA protein (Recombinase A).
GN Name=recA; OrderedLocuNames=plu1249;
OS Photorhabdus luminescens (subsp. laumondii).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Photorhabdus.
OX NCBI_TaxID=141679;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX MEDLINE=22957627; PubMed=14528314; DOI=10.1038/nbt886;
RA Duchaud E., Ruelink C., Frangeul L., Buchrieser C., Givaudan A.,
RA Taourit S., Bocs S., Boursaux-Eude C., Chandler M., Charles J.-P.,
RA Dassa E., Derose R., Derzelle S., Freysinet G., Gaudriault S.,
RA Medigue C., Lanois A., Powell K., Signier P., Vincent R., Wingate V.,
RA Zouine M., Glaser P., Boemare N., Danchin A., Kunst F.;
RT "The genome sequence of the entomopathogenic bacterium Photorhabdus
luminescens.";
RL Nat. Biotechnol. 21:1307-1313(2003).
CC -!- FUNCTION: Can catalyze the hydrolysis of ATP in the presence of
CC single-stranded DNA, the ATP-dependent uptake of single-stranded
CC DNA by duplex DNA, and the ATP-dependent hybridization of
CC homologous single-stranded DNAs. It interacts with lexA causing
CC its activation and leading to its autocatalytic cleavage (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the recA family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; BX571863; CAE13543.1; -; Genomic_DNA.
DR SMR; Q7N7A6; 4-329.
DR PhotoList; plu1249; -.
DR HAMAP; MF_002688; -; 1.
DR InterPro; IPR001553; RecA.
DR Pfam; PF00154; RecA; 1.
DR PRINTS; PR00142; RecA.
DR ProDom; PD000229; RecA; 1.
```

```
DR TIGRFAMs; TIGR02012; tigrfam_reca; 1.
DR PROSITE; PS00321; RECA_1; 1.
DR PROSITE; PS0162; RECA_2; 1.
DR PROSITE; PS0163; RECA_3; 1.
KW ATP-binding; Complete proteome; DNA damage; DNA recombination;
KW DNA repair; DNA-binding; Nucleotide-binding; SOS response.
FT NP_BIND 67 74 ATP (By similarity).
SQ SEQUENCE 355 AA; 38381 MW; PD659ACED827AE9 CRC64;

Query Match 92.2%; Score 1547; DB 1; Length 355;
Best Local Similarity 92.1%; Pred. No. 3.1e-95;
Matches 304; Conservative 15; Mismatches 11; Indels 0; Gaps 0;

QY 1 AIDENKOKALAAALGOIEKQFGKSGIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 60
Db 2 AIDENKOKALAAALGOIEKQFGKSGIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 61
QY 61 IVEIYGPSSGKTTTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 62 IVEIYGPSSGKTTTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 121
QY 121 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGEGSHMGLAARMMSQAMRKLKAG 180
Db 122 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGEGSHMGLAARMMSQAMRKLKAG 181
QY 181 NLKQNTLLIFINQIRMKIGVMFGNPTTTGGNALKFYASVRLDIRRIGAVKEGNNVVG 240
Db 182 NLKQNTLLIFINQIRMKIGVMFGNPTTTGGNALKFYASVRLDIRRIGAVKEGNNVVG 241
QY 241 ETRVVKVKNKIAAPFKQAEFQILYEGEGINFYGGELVDLGKVKELIEKAGAWSYNGEKIGQ 300
Db 242 ETRVVKVKNKIAAPFKQAEFQILYEGEGINFYGGELVDLGKVKELIEKAGAWSYNGEKIGQ 301
QY 301 GKANATWLKONPETAKIEKKVRELLLSN 330
Db 302 GKANATYLKEHPEVATELDKKREMLLHN 331

RESULT 11
RECA_ERWCT
ID RECA_ERWCT STANDARD; PRT; 357 AA.
AC Q6D1S8;
DT 10-MAY-2005 (Rel. 47, Created)
DT 10-MAY-2005 (Rel. 47, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE RecA protein (Recombinase A).
GN Name=recA; OrderedLocuNames=ECA3369;
OS Erwinia carotovora (subsp. atroseptica) (Pectobacterium atrosepticum).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Pectobacterium.
OX NCBI_TaxID=29471;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX STRAIN=SCRI 1043 / ATCC BAA-672;
RX PubMed=15263089; DOI=10.1073/pnas.0402424101;
RA Bell K.S., Sebahia M., Pritchard L., Holden M.T.G., Hyman L.J.,
RA Holleva M.C., Thomson N.R., Bentley S.D., Churcher I.J.C., Mungall K.,
RA Atkin K., Bason N., Brooks K., Chillingworth T., Clark K., Doggett J.,
RA Fraser A., Hance Z., Hauser H., Jagels K., Moule S., Norbertczak H.,
RA Ormond D., Price C., Quail M.A., Sanders M., Walker D., Whitehead S.,
RA Salmond G.P.C., Birch P.R.J., Parkhill J., Toth I.K.;
RT "Genome sequence of the enterobacterial phytopathogen Erwinia
carotovora subsp. atroseptica and characterization of virulence
factors.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:11105-11110(2004).
CC -!- FUNCTION: Can catalyze the hydrolysis of ATP in the presence of
CC single-stranded DNA, the ATP-dependent uptake of single-stranded
CC DNA by duplex DNA, and the ATP-dependent hybridization of
CC homologous single-stranded DNAs. It interacts with lexA causing
CC its activation and leading to its autocatalytic cleavage (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the recA family.
```

CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 DR EMBL; BX950851; CAG76267.1; -; Genomic_DNA.
 DR SMR; Q6D1S9; 4-329.
 DR HAMAP; MF_00268; -; 1.
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR001553; RecA.
 DR Pfam; PF00154; RecA; 1.
 DR PRINTS; PR00142; RecA.
 DR ProDom; PD000229; RecA; 1.
 DR SMART; SM00382; AAA; 1.
 DR TIGRFAMs; TIGR02012; tigrfam_recA; 1.
 DR PROSITE; PS00321; RecA_1; 1.
 DR PROSITE; PS50162; RecA_2; 1.
 DR PROSITE; PS50163; RecA_3; 1.
 KW ATP-binding; Complete proteome; DNA damage; DNA recombination;
 KW DNA repair; DNA-binding; Nucleotide-binding; SOS response.
 FT NP BIND 67 74 ATP (By similarity).
 SQ SEQUENCE 357 AA; 38284 MW; 48F1FFB8232E0098 CRC64;
 Query Match 91.6%; Score 1537; DB 1; Length 357;
 Best Local Similarity 91.6%; Pred. No. 1.4e-94;
 Matches 304; Conservative 14; Mismatches 14; Indels 0; Gaps 0;
 QY 1 AIDENKQKALAAALGOIEKQKQKSGIMRLGDRSDMDVTITGSLSLDIALGAGLPMGR 60
 DB 1 AIDENKQKALAAALGOIEKQKQKSGIMRLGDRSDMDVTITGSLSLDIALGAGLPMGR 61
 QY 61 IVEIYGPSSGKTYTLTQVIAAAREGKTCFIDAEHALDPIYAKLGVDDNLLCSQPD 120
 DB 62 IVEIYGPSSGKTYTLTQVIAAAREGKTCFIDAEHALDPIYAKLGVDDNLLCSQPD 121
 QY 121 TGEQALEICDALSAGVDVIVDSVAALTPKAEIEGIGDSHGLAARMWSQAMRKLKAG 180
 DB 122 TGEQALEICDALSAGVDVIVDSVAALTPKAEIEGIGDSHGLAARMWSQAMRKLKAG 181
 QY 181 NLKQSNLTLLIPIQIRKKGVMFGNPTTTCGNALKFYASVRLDIRRIGAVKEGVEVVG 240
 DB 182 NLKQANTLLIPIQIRKKGVMFGNPTTTCGNALKFYASVRLDIRRIGAVKEGVEVVG 241
 QY 241 ETRVKVKNKIAAPFKQAEFOILYGEINFGYELVDLGVKEKLEKAGAWYSYKGEIKGQ 300
 DB 242 ETRVKVKNKVAAPFKQAEFOILYGEINFGYELVDLGVKEKLEKAGAWYSYNGDKIKGQ 301
 QY 301 GKANATAWLKDNPTAKEIEKKVRELLISNPN 332
 DB 302 GKANACNFKENPTISAELDKKREMLLHKGN 333
 RESULT 12
 RECA YERPE
 ID RECA YERPE STANDARD; PRT; 355 AA.
 AC P37858;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE RecA protein (Recombinase A).
 GN Name=recA; OrderedLocNames=YPO3307, Y0881, YP0379;
 OS Versinia pestis.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Versinia.
 OX NCBI_TaxID=632;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=231;
 RA Kryukov V.M., Suchkov I.Y., Sazykin I.S., Mishankin B.N.;
 RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.
 RN [2]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=CO-92 / Biovar Orientalis;
 RX MEDLINE=21470413; PubMed=11586360; DOI=10.1038/35097083;
 RA Parhill J., Wren B.W., Thomson N.R., Tibball R.W., Holden M.T.G.,
 RA Pringle M.B., Sebahia M., James K.D., Churcher C.M., Mungall K.L.,
 RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Rarraga A.-M.,
 RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
 RA Fellwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
 RA Leather S., Moule S., Oyston P.C.F., Quail M.A., Rutherford K.M.,
 RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
 RT "Genome sequence of Versinia pestis, the causative agent of plague.";
 RL Nature 413:523-527(2001).
 RN [3]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=KIM5 / Biovar Mediaevalis;
 RX MEDLINE=22137863; PubMed=12142430;
 RX DOI=10.1128/JB.184.16.4601-4611.2002;
 RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
 RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
 RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
 RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
 RA Perry R.D.;
 RT "Genome sequence of Versinia pestis KIM.";
 RL J. Bacteriol. 184:4601-4611(2002).
 RN [4]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=91001 / Biovar Mediaevalis;
 RX PubMed=15368893;
 RA Song Y., Tong Z., Wang J., Wang L., Guo Z., Han Y., Zhang J., Pei D.,
 RA Zhou D., Qin H., Pang X., Han Y., Zhai J., Li M., Cui B., Qi Z.,
 RA Jin L., Dai R., Chen F., Li S., Ye C., Du Z., Lin W., Wang J., Yu J.,
 RA Yang H., Wang J., Huang P., Yang R.;
 RT "Complete genome sequence of Versinia pestis strain 91001, an isolate
 avirulent to humans.";
 RL DNA Res. 11:179-197(2004).
 CC -!- FUNCTION: Can catalyze the hydrolysis of ATP in the presence of
 CC single-stranded DNA, the ATP-dependent uptake of single-stranded
 CC DNA by duplex DNA, and the ATP-dependent hybridization of
 CC homologous single-stranded DNAs. It interacts with LexA causing
 CC its activation and leading to its autocatalytic cleavage.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- SIMILARITY: Belongs to the recA family.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 DR EMBL; X75336; CAA53084.1; -; Genomic DNA.
 DR EMBL; AJ414156; CAC92539.1; -; Genomic DNA.
 DR EMBL; AE013691; AAM84465.1; -; Genomic DNA.
 DR EMBL; AE017128; AAS60652.1; -; Genomic DNA.
 DR PIR; AG0401; AG0401.
 DR PIR; S37586; S37586.
 DR HSPP; P03017; 2REB.
 DR SMR; P37858; 3-328.
 DR HAMAP; MF_00268; -; 1.
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR001553; RecA.
 DR Pfam; PF00154; RecA; 1.
 DR PRINTS; PR00142; RecA.
 DR ProDom; PD000229; RecA; 1.
 DR SMART; SM00382; AAA; 1.
 DR TIGRFAMs; TIGR02012; tigrfam_recA; 1.
 DR PROSITE; PS00321; RecA_1; 1.
 DR PROSITE; PS50162; RecA_2; 1.
 DR PROSITE; PS50163; RecA_3; 1.
 KW ATP-binding; Complete proteome; DNA damage; DNA recombination;
 KW DNA repair; DNA-binding; Nucleotide-binding; SOS response.
 FT INIT MET 0 0 By similarity.
 FT NP BIND 66 73 ATP (By similarity).
 FT CONFLICT 281 281 H -> L (in Ref. 1).

FT CONFLICT 294 294 N -> Y (in Ref. 1).
 FT CONFLICT 315 315 I -> N (in Ref. 1).
 SQ SEQUENCE 355 AA; 37755 MW; F800D2D1AD32AB81 CRC64;
 Query Match 91.4%; Score 1534; DB 1; Length 355;
 Best Local Similarity 91.9%; Pred. No. 2.3e-94;
 Matches 305; Conservative 14; Mismatches 13; Indels 0; Gaps 0;
 QY 1 AIDENKOKALAAALGOIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 60
 Db 1 AIDENKOKALAAALGOIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 60
 QY 61 IVEIYGPSSSGKTTLTQVIAAAQREGKTCFADAEHALDPIYAKKLGVDIDNLLCSQPD 120
 Db 61 IVEIYGPSSSGKTTLTQVIAAAQREGKTCFADAEHALDPIYAKKLGVDIDNLLCSQPD 120
 QY 121 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKLAG 180
 Db 121 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKLAG 180
 QY 181 NLKQNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
 Db 181 NLKQNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
 QY 241 ETRVVKVKNKIAAPFKQAEFOILYEGEINFGELVDLGKVKELIEKAGAWYSYKGEKIQ 300
 Db 241 ETRVVKVKNKIAAPFKQAEFOILYEGEINFGELVDLGKVKELIEKAGAWYSYKGEKIQ 300
 QY 301 GKANATWLKNDPETAKEIEKKVRELLLSNP 332
 Db 301 GKANASNYLKENPAIAELDKKREMLLNGN 332

RESULT 13

RECA YERPS STANDARD; PRT; 356 AA.
 ID RECA_YERPS STANDARD; PRT; 356 AA.
 AC Q66E70;
 DT 10-MAY-2005 (Rel. 47, Created)
 DT 10-MAY-2005 (Rel. 47, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE RecA protein (Recombinase A).
 GN Name=recA; OrderedLocusName=YPTB0823;
 OS Yersinia pseudotuberculosis.
 OC Bacteria; Proteobacteria;
 OC Enterobacteriaceae; Yersinia.
 ON NCBI_TaxID=633;
 RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=IP32953 / Serotype I;
 RX PubMed=15358858; DOI=10.1073/pnas.040401201;
 RA Chain P.S.G., Carniel E., Larimer F.W., Lamerdin J., Stoutland P.O.,
 RA Regala W.M., Georgescu A.M., Verges L.M., Land M.L., Motin V.L.,
 RA Brubaker R.R., Fowler J., Hinnebusch J., Marceau M., Medigue C.,
 RA Simonet M., Chenal-Francois V., Souza B., Dacheux D., Elliott J.M.,
 RA Derbise A., Hauser L.J., Garcia E.;
 RT "Insights into the evolution of Yersinia pestis through whole-genome
 comparison with Yersinia pseudotuberculosis.";
 RL Proc. Natl. Acad. Sci. U.S.A. 101:13826-13831(2004).
 CC -!- FUNCTION: Can catalyze the hydrolysis of ATP in the presence of
 CC single-stranded DNA, and the ATP-dependent uptake of single-stranded
 CC DNA by duplex DNA, and the ATP-dependent hybridization of
 CC homologous single-stranded DNAs. It interacts with LexA causing
 CC its activation and leading to its autocatalytic cleavage (By
 CC similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- SIMILARITY: Belongs to the RecA family.
 CC
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.

DR EMBL; BX936398; CAH20063.1; -, Genomic_DNA.
 DR SMR; Q66E70; 4-329.
 DR HAWAP; MF_00268; -, 1.
 DR InterPro; IPR003593; AAA_ATPase.
 DR Pfam; PF001553; RecA.
 DR Pfam; PF00154; RecA; 1.
 DR PRINTS; PR00142; RECA.
 DR ProDom; PD000229; RecA; 1.
 DR SMART; SMO0382; AAA; 1.
 DR TIGRFAMs; TIGR02012; tigrfam_recA; 1.
 DR PROSITE; PS00321; RECA_1; 1.
 DR PROSITE; PS0162; RECA_2; 1.
 DR PROSITE; PS0163; RECA_3; 1.
 KW ATP-binding; Complete proteome; DNA damage; DNA recombination;
 KW DNA repair; DNA-binding; Nucleotide-binding; SOS response.
 FT NP BIND 67 74 ATP (By similarity).
 SQ SEQUENCE 356 AA; 37914 MW; DIC0D90F58C642B CRC64;
 Query Match 91.4%; Score 1534; DB 1; Length 356;
 Best Local Similarity 91.9%; Pred. No. 2.3e-94;
 Matches 305; Conservative 14; Mismatches 13; Indels 0; Gaps 0;
 QY 1 AIDENKOKALAAALGOIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 60
 Db 2 AIDENKOKALAAALGOIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 61
 QY 61 IVEIYGPSSSGKTTLTQVIAAAQREGKTCFADAEHALDPIYAKKLGVDIDNLLCSQPD 120
 Db 62 IVEIYGPSSSGKTTLTQVIAAAQREGKTCFADAEHALDPIYAKKLGVDIDNLLCSQPD 121
 QY 121 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKLAG 180
 Db 122 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKLAG 181
 QY 181 NLKQNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
 Db 182 NLKQNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 241
 QY 241 ETRVVKVKNKIAAPFKQAEFOILYEGEINFGELVDLGKVKELIEKAGAWYSYKGEKIQ 300
 Db 242 ETRVVKVKNKIAAPFKQAEFOILYEGEINFGELVDLGKVKELIEKAGAWYSYKGEKIQ 301
 QY 301 GKANATWLKNDPETAKEIEKKVRELLLSNP 332
 Db 302 GKANASNYLKENPAIAELDKKREMLLNGN 333
 RESULT 14
 RECA_ENTAG STANDARD; PRT; 353 AA.
 ID RECA_ENTAG STANDARD; PRT; 353 AA.
 AC P33037;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE RecA protein (Recombinase A).
 GN Name=recA;
 OS Enterobacter agglomerans (Erwinia herbicola) (Pantoea agglomerans).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Pantoea.
 ON NCBI_TaxID=549;
 RN NUCLEOTIDE SEQUENCE [GENOMIC DNA].
 RP Rappold C.S.J., Klingmueller W.;
 RL Submitted (JAN-1993) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Can catalyze the hydrolysis of ATP in the presence of
 CC single-stranded DNA, the ATP-dependent uptake of single-stranded
 CC DNA by duplex DNA, and the ATP-dependent hybridization of
 CC homologous single-stranded DNAs. It interacts with LexA causing
 CC its activation and leading to its autocatalytic cleavage.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- SIMILARITY: Belongs to the RecA family.
 CC
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.

DR EMBL; L03291; AAA91766.1; -; Genomic_DNA.
DR PIR; S1481; S31481.
DR HSSP; P03017; 2REB.
DR SMR; P33037; 3-328.
DR HAMAP; MF 00268; -; 1.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001553; RecA.
DR Pfam; PF00154; RecA; 1.
DR PRINTS; PR00142; RecA; 1.
DR ProDom; PD000229; RecA; 1.
DR SMART; SM00382; AAA; 1.
DR TIGRFAMs; TIGR02012; tigrfam_recA; 1.
DR PROSITE; PS00321; RECA_1; 1.
DR PROSITE; PS50162; RECA_2; 1.
DR PROSITE; PS50163; RECA_3; 1.
KW ATP-binding; DNA damage; DNA recombination; DNA repair; DNA-binding;
KW Nucleotide-binding; SOS response.
FT INIT MET 0 By similarity.
FT NP_BIND 66 73 ATP (By similarity).
SQ SEQUENCE 353 AA; 37767 MW; 960F1F2698CAC46B CRC64;

Query Match 91.3%; Score 1532; DB 1; Length 353;
Best Local Similarity 92.1%; Pred. No. 3.1e-94;
Matches 303; Conservative 15; Mismatches 11; Indels 0; Gaps 0;

Qy 1 AIDENKQKALAAALGQIEKQFGKSGIMRLGEDRSMVDVETISTGSLDIALGAGGLPMGR 60
Db 1 AIDENKQKALAAALGQIEKQFGKSGIMRLGEDRSMVDVETISTGSLDIALGAGGLPMGR 60
Qy 61 IVEIYGPESSGKTTTLQVIAAAQREGKTCAFIDAEHALDPYAKGLGVDIDNLLCSQPD 120
Db 61 IVEIYGPESSGKTTTLQVIAAAQREGKTCAFIDAEHALDPYAKGLGVDIDNLLCSQPD 120
Qy 121 TGEQALEICDALARSGADVIVVDSVAALTPKAEIEGEGISHMGLAARMMSQAMRKLKAG 180
Db 121 TGEQALEICDALARSGADVIVVDSVAALTPKAEIEGEGISHMGLAARMMSQAMRKLKAG 180
Qy 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
Db 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
Qy 241 ETRVKVKNKTAAPPKQAEFOILYGEINFGELVDLGKVEKLEKAGAWYSYKGEKITGQ 300
Db 241 ETRVKVKNKTAAPPKQAEFOILYGEINFGELVDLGKVEKLEKAGAWYSYKGEKITGQ 300
Qy 301 GKANATAWLKDNPETAKEIEKKVRELLLS 329
Db 301 GKANSCNYLKENPKVAELDKKLRDMLLS 329

RESULT 15

RECA_SERMA STANDARD; PRT; 353 AA.
AC P17479;
DT 01-AUG-1990 (Rel. 15, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE RecA protein (Recombinase A).
GN Name=recA;
OS Serratia marcescens.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Serratia.
OX NCBI_TaxID=615;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RC STRAIN=SM6;
RX MEDLINE=90094239; PubMed=2152908;
RA Ball T.K., Wasmuth C.R., Braunagel S.C., Benedik M.J.;

RT "Expression of Serratia marcescens extracellular proteins requires
RT recA".
RL J. Bacteriol. 172:342-349(1990).

CC -!- FUNCTION: Can catalyze the hydrolysis of ATP in the presence of
CC single-stranded DNA, and the ATP-dependent uptake of single-stranded
CC DNA by duplex DNA, and the ATP-dependent hybridization of
CC homologous single-stranded DNAs. It interacts with lexA causing
CC its activation and leading to its autocatalytic cleavage.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the recA family.

CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.

CC EMBL; M22935; AAA26567.1; -; Genomic_DNA.
CC HSSP; P03017; 2REB.
CC SMR; P17479; 3-328.
CC HAMAP; MF 00268; -; 1.
CC InterPro; IPR003593; AAA_ATPase.
CC InterPro; IPR001553; RecA.
CC Pfam; PF00154; RecA; 1.
CC PRINTS; PR00142; RecA; 1.
CC ProDom; PD000229; RecA; 1.
CC SMART; SM00382; AAA; 1.
CC TIGRFAMs; TIGR02012; tigrfam_recA; 1.
CC PROSITE; PS00321; RECA_1; 1.
CC PROSITE; PS50162; RECA_2; 1.
CC PROSITE; PS50163; RECA_3; 1.
KW ATP-binding; DNA damage; DNA recombination; DNA repair; DNA-binding;
KW Nucleotide-binding; SOS response.
FT INIT MET 0 By similarity.
FT NP_BIND 66 73 ATP (By similarity).
SQ SEQUENCE 353 AA; 37778 MW; A97369360970F814 CRC64;

Query Match 91.0%; Score 1527; DB 1; Length 353;
Best Local Similarity 92.4%; Pred. No. 6.6e-94;
Matches 303; Conservative 13; Mismatches 12; Indels 0; Gaps 0;

Qy 1 AIDENKQKALAAALGQIEKQFGKSGIMRLGEDRSMVDVETISTGSLDIALGAGGLPMGR 60
Db 1 AIDENKQKALAAALGQIEKQFGKSGIMRLGEDRSMVDVETISTGSLDIALGAGGLPMGR 60
Qy 61 IVEIYGPESSGKTTTLQVIAAAQREGKTCAFIDAEHALDPYAKGLGVDIDNLLCSQPD 120
Db 61 IVEIYGPESSGKTTTLQVIAAAQREGKTCAFIDAEHALDPYAKGLGVDIDNLLCSQPD 120
Qy 121 TGEQALEICDALARSGADVIVVDSVAALTPKAEIEGEGISHMGLAARMMSQAMRKLKAG 180
Db 121 TGEQALEICDALARSGADVIVVDSVAALTPKAEIEGEGISHMGLAARMMSQAMRKLKAG 180
Qy 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
Db 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
Qy 241 ETRVKVKNKTAAPPKQAEFOILYGEINFGELVDLGKVEKLEKAGAWYSYKGEKITGQ 300
Db 241 ETRVKVKNKTAAPPKQAEFOILYGEINFGELVDLGKVEKLEKAGAWYSYKGEKITGQ 300
Qy 301 GKANATAWLKDNPETAKEIEKKVRELLL 328
Db 301 GKANACFLKENPAIAAELEKLRDILL 328

Search completed: February 16, 2006, 01:24:00
Job time : 153 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 16, 2006, 01:24:19 ; Search time 32.5 Seconds
(without alignments)
852.195 Million cell updates/sec

Title: US-10-733-782-1
Perfect score: 1678
Sequence: 1 AIDENKQKALAAALGQIEKQ.....AKIEKKVRELLSNPNSTP 335

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/1/iaa/5 COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/6 COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/H COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/PCUTS COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/RE COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1678	100.0	353	1	US-08-097-831-6
2	1678	100.0	358	2	US-09-626-410-7
3	1678	100.0	358	2	US-09-626-410-9
4	1678	100.0	358	2	US-09-626-410-14
5	1678	100.0	358	2	US-09-116-188-7
6	1678	100.0	358	2	US-09-116-188-9
7	1678	100.0	358	2	US-09-116-188-14
8	1678	100.0	358	2	US-09-626-047-9
9	1678	100.0	358	2	US-09-626-047-9
10	1678	100.0	358	2	US-09-626-047-14
11	1678	100.0	358	2	US-09-626-343-7
12	1678	100.0	358	2	US-09-626-343-9
13	1678	100.0	358	2	US-09-626-343-14
14	1678	100.0	358	2	US-09-354-922-8
15	1678	100.0	358	2	US-09-354-922-10
16	1678	100.0	358	2	US-09-354-922-15
17	1678	100.0	358	2	US-09-516-051-7
18	1678	100.0	358	2	US-09-516-051-9
19	1678	100.0	358	2	US-09-516-051-14
20	1678	100.0	358	2	US-09-516-695B-8
21	1678	100.0	358	2	US-09-516-695B-10
22	1678	100.0	358	2	US-09-516-695B-15
23	1673	99.7	358	2	US-09-626-410-8
24	1673	99.7	358	2	US-09-626-410-11
25	1673	99.7	358	2	US-09-116-188-8
26	1673	99.7	358	2	US-09-116-188-11
27	1673	99.7	358	2	US-09-626-047-8

28	1673	99.7	358	2	US-09-626-047-11	Sequence 11, Appl
29	1673	99.7	358	2	US-09-626-343-8	Sequence 8, Appl
30	1673	99.7	358	2	US-09-626-343-11	Sequence 11, Appl
31	1673	99.7	358	2	US-09-354-922-9	Sequence 9, Appl
32	1673	99.7	358	2	US-09-354-922-12	Sequence 12, Appl
33	1673	99.7	358	2	US-09-516-051-8	Sequence 8, Appl
34	1673	99.7	358	2	US-09-516-051-11	Sequence 11, Appl
35	1673	99.7	358	2	US-09-516-695B-9	Sequence 9, Appl
36	1673	99.7	358	2	US-09-516-695B-12	Sequence 12, Appl
37	1660	98.9	358	2	US-09-626-410-12	Sequence 12, Appl
38	1660	98.9	358	2	US-09-116-188-12	Sequence 12, Appl
39	1660	98.9	358	2	US-09-626-047-12	Sequence 12, Appl
40	1660	98.9	358	2	US-09-626-343-12	Sequence 12, Appl
41	1660	98.9	358	2	US-09-354-922-13	Sequence 13, Appl
42	1660	98.9	358	2	US-09-516-051-12	Sequence 12, Appl
43	1660	98.9	358	2	US-09-516-695B-13	Sequence 13, Appl
44	1659	98.9	358	2	US-09-626-410-10	Sequence 10, Appl
45	1659	98.9	358	2	US-09-116-188-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-08-097-831-6
; Sequence 6, Application US/08097831
; Patent No. 5510473
; GENERAL INFORMATION:
; APPLICANT: Camerini-Otero, Rafael D.
; APPLICANT: Angov, Evangelina
; TITLE OF INVENTION: Cloning and Expression of Taq recA
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KNOBBE, MARTENS, OLSON AND BEAR
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: USA
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/097,831
; APPLICATION NUMBER: US/08/097,831
; FILING DATE: 19930726
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Altman, Daniel E.
; REGISTRATION NUMBER: 34,115
; REFERENCE/DOCKET NUMBER: NIH066.001A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 714-760-0404
; TELEFAX: 714-760-9502
; INFORMATION FOR SEQ ID NO: 6:
; LENGTH: 353 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
US-08-097-831-6

Query Match 100.0%; Score 1678; DB 1; Length 353;
Best Local Similarity 100.0%; Pred. No. 2.5e-167;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AIDENKQKALAAALGQIEKQIMRLGEDSRMDVETISTGSLSLDIALGAGLPMGR 60
|||||

Db 2 AIDENKQKALAAALGQIEKQFGKSGIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 61
 Qy 61 IVEIYGPSSGKTTTLTQVIAAAQREGKTCFADAEHALDPIYARKLGVDIDNLLCSQPD 120
 Db 62 IVEIYGPSSGKTTTLTQVIAAAQREGKTCFADAEHALDPIYARKLGVDIDNLLCSQPD 121
 Qy 121 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKLKAG 180
 Db 122 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKLKAG 181
 Qy 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVG 240
 Db 182 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVG 241
 Qy 241 ETRVVKVKNKIAAPFKQAEFQILYGEINFGYELVDLGVKEKLIIEKAGAWYSYKGEKIGQ 300
 Db 242 ETRVVKVKNKIAAPFKQAEFQILYGEINFGYELVDLGVKEKLIIEKAGAWYSYKGEKIGQ 301
 Qy 301 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 335
 Db 302 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 336

RESULT 2
 US-09-626-410-7
 ; Sequence 7, Application US/09626410
 ; Patent No. 6287862
 ; GENERAL INFORMATION:
 ; APPLICANT: DEL CARDAYRE, STEPHEN
 ; APPLICANT: TOBIN, MATTHEW
 ; APPLICANT: STEMMER, WILLEM P.C.
 ; APPLICANT: NESS, JON E.
 ; APPLICANT: MINSHULL, JEREMY
 ; APPLICANT: PATTEN, PHILLIP
 ; APPLICANT: SUBRAMANIA, VENKITSWARAN
 ; APPLICANT: CASTLE, LINDA
 ; APPLICANT: KREBBER, CLAUD M.
 ; APPLICANT: BASS, STEVE
 ; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
 ; FILE REFERENCE: 02-020720US
 ; CURRENT APPLICATION NUMBER: US/09/626,410
 ; PRIOR FILING DATE: 2000-07-26
 ; PRIOR APPLICATION NUMBER: 09/116,188
 ; PRIOR FILING DATE: 07-15-1998
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: Patentin Ver. 2.0
 ; SEQ ID NO 7
 ; LENGTH: 358
 ; TYPE: PRT
 ; ORGANISM: Escherichia coli
 US-09-626-410-7

Query Match 100.0%; Score 1678; DB 2; Length 358;
 Best Local Similarity 100.0%; Pred. No. 2.6e-167;
 Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AIDENKQKALAAALGQIEKQFGKSGIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 60
 Db 7 AIDENKQKALAAALGQIEKQFGKSGIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 66
 Qy 61 IVEIYGPSSGKTTTLTQVIAAAQREGKTCFADAEHALDPIYARKLGVDIDNLLCSQPD 120
 Db 67 IVEIYGPSSGKTTTLTQVIAAAQREGKTCFADAEHALDPIYARKLGVDIDNLLCSQPD 126
 Qy 121 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKLKAG 180
 Db 127 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKLKAG 186
 Qy 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVG 240
 Db 187 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVG 246

Qy 241 ETRVVKVKNKIAAPFKQAEFQILYGEINFGYELVDLGVKEKLIIEKAGAWYSYKGEKIGQ 300
 Db 247 ETRVVKVKNKIAAPFKQAEFQILYGEINFGYELVDLGVKEKLIIEKAGAWYSYKGEKIGQ 306
 Qy 301 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 335
 Db 307 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 341

RESULT 3
 US-09-626-410-9
 ; Sequence 9, Application US/09626410
 ; Patent No. 6287862
 ; GENERAL INFORMATION:
 ; APPLICANT: DEL CARDAYRE, STEPHEN
 ; APPLICANT: TOBIN, MATTHEW
 ; APPLICANT: STEMMER, WILLEM P.C.
 ; APPLICANT: NESS, JON E.
 ; APPLICANT: MINSHULL, JEREMY
 ; APPLICANT: PATTEN, PHILLIP
 ; APPLICANT: SUBRAMANIA, VENKITSWARAN
 ; APPLICANT: CASTLE, LINDA
 ; APPLICANT: KREBBER, CLAUD M.
 ; APPLICANT: BASS, STEVE
 ; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
 ; FILE REFERENCE: 02-020720US
 ; CURRENT APPLICATION NUMBER: US/09/626,410
 ; CURRENT FILING DATE: 2000-07-26
 ; PRIOR APPLICATION NUMBER: 09/116,188
 ; PRIOR FILING DATE: 07-15-1998
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: Patentin Ver. 2.0
 ; SEQ ID NO 9
 ; LENGTH: 358
 ; TYPE: PRT
 ; ORGANISM: Escherichia coli
 US-09-626-410-9

Query Match 100.0%; Score 1678; DB 2; Length 358;
 Best Local Similarity 100.0%; Pred. No. 2.6e-167;
 Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AIDENKQKALAAALGQIEKQFGKSGIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 60
 Db 7 AIDENKQKALAAALGQIEKQFGKSGIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 66
 Qy 61 IVEIYGPSSGKTTTLTQVIAAAQREGKTCFADAEHALDPIYARKLGVDIDNLLCSQPD 120
 Db 67 IVEIYGPSSGKTTTLTQVIAAAQREGKTCFADAEHALDPIYARKLGVDIDNLLCSQPD 126
 Qy 121 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKLKAG 180
 Db 127 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKLKAG 186
 Qy 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVG 240
 Db 187 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVG 246
 Qy 241 ETRVVKVKNKIAAPFKQAEFQILYGEINFGYELVDLGVKEKLIIEKAGAWYSYKGEKIGQ 300
 Db 247 ETRVVKVKNKIAAPFKQAEFQILYGEINFGYELVDLGVKEKLIIEKAGAWYSYKGEKIGQ 306
 Qy 301 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 335
 Db 307 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 341

RESULT 4
 US-09-626-410-14
 ; Sequence 14, Application US/09626410
 ; Patent No. 6287862
 ; GENERAL INFORMATION:

; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIA, VENKITSWARAN
; APPLICANT: CASTLE, LINDA
; APPLICANT: KREBBER, CLAUD M.
; APPLICANT: BASS, STEVE
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
; FILE REFERENCE: 02-02072005
; CURRENT APPLICATION NUMBER: US/09/626,410
; CURRENT FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 09/116,188
; PRIOR FILING DATE: 07-15-1998
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: consensus
; OTHER INFORMATION: e. coli sequence
US-09-626-410-14

Query Match 100.0%; Score 1678; DB 2; Length 358;
Best Local Similarity 100.0%; Pred. No. 2.6e-167;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AIDENKQKALAAALGQIEKQFGKSGIMRLGDRSDMDVETISTGSLSLDIALGAGGLPMGR 60
Db 7 AIDENKQKALAAALGQIEKQFGKSGIMRLGDRSDMDVETISTGSLSLDIALGAGGLPMGR 66

Qy 61 IVEIYGPESSGKTTLTLOVIAAAREGKTCAPIDAEHALDPIYARKLGVDINDLLCSQPD 120
Db 67 IVEIYGPESSGKTTLTLOVIAAAREGKTCAPIDAEHALDPIYARKLGVDINDLLCSQPD 126

Qy 121 TGEQALEICDALARSGAVDVIVDSVAALTPKAEIEGIGSHMGLAARMMSQAMRKLKAG 180
Db 127 TGEQALEICDALARSGAVDVIVDSVAALTPKAEIEGIGSHMGLAARMMSQAMRKLKAG 186

Qy 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGNNALKFYASVRLDIRRIGAVKEGNNVGS 240
Db 187 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGNNALKFYASVRLDIRRIGAVKEGNNVGS 246

Qy 241 ETRVKVVKNTAAAPFKQAEFQILYGEINFGELVDLGKVKLEKAGAWSYKGEKIGQ 300
Db 247 ETRVKVVKNTAAAPFKQAEFQILYGEINFGELVDLGKVKLEKAGAWSYKGEKIGQ 306

Qy 301 GKANATLWKNPETAKEIEKKVRELLLSNPNSTP 335
Db 307 GKANATLWKNPETAKEIEKKVRELLLSNPNSTP 341

RESULT 5
US-09-116-188-7
; Sequence 7, Application US/09116188
; Patent No. 6326204
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIA, VENKITSWARAN
; APPLICANT: CASTLE, LINDA
; APPLICANT: KREBBER, CLAUD M.
; APPLICANT: BASS, STEVE
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE

; TITLE OF INVENTION: SEQUENCE RECOMBINATION
; FILE REFERENCE: 02-02072005
; CURRENT APPLICATION NUMBER: US/09/116,188
; CURRENT FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-116-188-7

Query Match 100.0%; Score 1678; DB 2; Length 358;
Best Local Similarity 100.0%; Pred. No. 2.6e-167;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AIDENKQKALAAALGQIEKQFGKSGIMRLGDRSDMDVETISTGSLSLDIALGAGGLPMGR 60
Db 7 AIDENKQKALAAALGQIEKQFGKSGIMRLGDRSDMDVETISTGSLSLDIALGAGGLPMGR 66

Qy 61 IVEIYGPESSGKTTLTLOVIAAAREGKTCAPIDAEHALDPIYARKLGVDINDLLCSQPD 120
Db 67 IVEIYGPESSGKTTLTLOVIAAAREGKTCAPIDAEHALDPIYARKLGVDINDLLCSQPD 126

Qy 121 TGEQALEICDALARSGAVDVIVDSVAALTPKAEIEGIGSHMGLAARMMSQAMRKLKAG 180
Db 127 TGEQALEICDALARSGAVDVIVDSVAALTPKAEIEGIGSHMGLAARMMSQAMRKLKAG 186

Qy 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGNNALKFYASVRLDIRRIGAVKEGNNVGS 240
Db 187 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGNNALKFYASVRLDIRRIGAVKEGNNVGS 246

Qy 241 ETRVKVVKNTAAAPFKQAEFQILYGEINFGELVDLGKVKLEKAGAWSYKGEKIGQ 300
Db 247 ETRVKVVKNTAAAPFKQAEFQILYGEINFGELVDLGKVKLEKAGAWSYKGEKIGQ 306

Qy 301 GKANATLWKNPETAKEIEKKVRELLLSNPNSTP 335
Db 307 GKANATLWKNPETAKEIEKKVRELLLSNPNSTP 341

RESULT 6
US-09-116-188-9
; Sequence 9, Application US/09116188
; Patent No. 6326204
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIA, VENKITSWARAN
; APPLICANT: CASTLE, LINDA
; APPLICANT: KREBBER, CLAUD M.
; APPLICANT: BASS, STEVE
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
; FILE REFERENCE: 02-02072005
; CURRENT APPLICATION NUMBER: US/09/116,188
; CURRENT FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-116-188-9

Query Match 100.0%; Score 1678; DB 2; Length 358;
Best Local Similarity 100.0%; Pred. No. 2.6e-167;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 60
DB 7 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 66
QY 61 IVEIYGPESSGKTTTLQVIAAAQREGKTCAPIDAHAHALDPIYARKLGVDIDNLLCSQPD 120
DB 67 IVEIYGPESSGKTTTLQVIAAAQREGKTCAPIDAHAHALDPIYARKLGVDIDNLLCSQPD 126
QY 121 TGEQALEICDALARSGAVDVIIVDVSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKLKAG 180
DB 127 TGEQALEICDALARSGAVDVIIVDVSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKLKAG 186
QY 181 NLKQSNNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENNVVGS 240
DB 187 NLKQSNNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENNVVGS 246
QY 241 ETRVKVKNKIAAPFKQAEFQILYGEINFGYELVDLGKVEKLEKAGAWSYKGEKIGQ 300
DB 247 ETRVKVKNKIAAPFKQAEFQILYGEINFGYELVDLGKVEKLEKAGAWSYKGEKIGQ 306
QY 301 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 335
DB 307 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 341
RESULT 7
US-09-116-188-14
; Sequence 14, Application US/09116188
; Patent No. 6326204
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIA, VENKITSWARAN
; APPLICANT: CASTLE, LINDA
; APPLICANT: KREBBER, CLAUD M.
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
; FILE REFERENCE: 02-020720US
; CURRENT FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: US/09/116,188
; CURRENT FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: consensus
US-09-116-188-14
Query Match 100.0%; Score 1678; DB 2; Length 358;
Best Local Similarity 100.0%; Pred. No. 2.6e-167;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 60
DB 7 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 66
QY 61 IVEIYGPESSGKTTTLQVIAAAQREGKTCAPIDAHAHALDPIYARKLGVDIDNLLCSQPD 120
DB 67 IVEIYGPESSGKTTTLQVIAAAQREGKTCAPIDAHAHALDPIYARKLGVDIDNLLCSQPD 126
QY 121 TGEQALEICDALARSGAVDVIIVDVSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKLKAG 180
DB 127 TGEQALEICDALARSGAVDVIIVDVSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKLKAG 186
QY 181 NLKQSNNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENNVVGS 240
```

```
DB 187 NLKQSNNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENNVVGS 246
QY 241 ETRVKVKNKIAAPFKQAEFQILYGEINFGYELVDLGKVEKLEKAGAWSYKGEKIGQ 300
DB 247 ETRVKVKNKIAAPFKQAEFQILYGEINFGYELVDLGKVEKLEKAGAWSYKGEKIGQ 306
QY 301 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 335
DB 307 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 341
RESULT 8
US-09-626-047-7
; Sequence 7, Application US/09626047
; Patent No. 6335198
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIA, VENKITSWARAN
; APPLICANT: CASTLE, LINDA
; APPLICANT: KREBBER, CLAUD M.
; APPLICANT: BASS, STEVE
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
; FILE REFERENCE: 02-020720US
; CURRENT APPLICATION NUMBER: US/09/626,047
; CURRENT FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 09/116,188
; PRIOR FILING DATE: 07-15-1998
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-626-047-7
Query Match 100.0%; Score 1678; DB 2; Length 358;
Best Local Similarity 100.0%; Pred. No. 2.6e-167;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 60
DB 7 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 66
QY 61 IVEIYGPESSGKTTTLQVIAAAQREGKTCAPIDAHAHALDPIYARKLGVDIDNLLCSQPD 120
DB 67 IVEIYGPESSGKTTTLQVIAAAQREGKTCAPIDAHAHALDPIYARKLGVDIDNLLCSQPD 126
QY 121 TGEQALEICDALARSGAVDVIIVDVSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKLKAG 180
DB 127 TGEQALEICDALARSGAVDVIIVDVSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKLKAG 186
QY 181 NLKQSNNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENNVVGS 240
DB 187 NLKQSNNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENNVVGS 246
QY 241 ETRVKVKNKIAAPFKQAEFQILYGEINFGYELVDLGKVEKLEKAGAWSYKGEKIGQ 300
DB 247 ETRVKVKNKIAAPFKQAEFQILYGEINFGYELVDLGKVEKLEKAGAWSYKGEKIGQ 306
QY 301 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 335
DB 307 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 341
RESULT 9
US-09-626-047-9
```

; Sequence 9, Application US/09626047
; Patent No. 6335198
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIA, VENKITSWARAN
; APPLICANT: CASTLE, LINDA
; APPLICANT: KREBER, CLAUD M.
; APPLICANT: BASS, STEVE
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
; FILE REFERENCE: 02-020720US
; CURRENT APPLICATION NUMBER: US/09/626,047
; CURRENT FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 09/116,188
; PRIOR FILING DATE: 07-15-1998
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-626-047-9

Query Match 100.0%; Score 1678; DB 2; Length 358;
Best Local Similarity 100.0%; Pred. No. 2.6e-167;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AIDENKQKALAAALGQIEKQFGKSGIMRLGEDRSMDVETISTGSLSLDIALGAGGLPMGR 60
DB 7 AIDENKQKALAAALGQIEKQFGKSGIMRLGEDRSMDVETISTGSLSLDIALGAGGLPMGR 66
QY 61 IVEIYGPESSGKTTTLTQVIAAAQREGKTCAPIDAEHALDPIYARKLGVDIDNLLCSQPD 120
DB 67 IVEIYGPESSGKTTTLTQVIAAAQREGKTCAPIDAEHALDPIYARKLGVDIDNLLCSQPD 126
QY 121 TGEQALEICDALARGAVDVIIVDSVAALTPKAEIEGEIGDISHMGLAARMMSQAMRKLKAG 180
DB 127 TGEQALEICDALARGAVDVIIVDSVAALTPKAEIEGEIGDISHMGLAARMMSQAMRKLKAG 186
QY 181 NLKQSNLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVG 240
DB 187 NLKQSNLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVG 246
QY 241 ETRVVKVKNKIAAPPKQAEFQILYGEINFGYELVDLGKELIEKAGAWYSYKGEKIQ 300
DB 247 ETRVVKVKNKIAAPPKQAEFQILYGEINFGYELVDLGKELIEKAGAWYSYKGEKIQ 306
QY 301 GKANATAMKONPETAKEIEKKVRELLLSNPSTP 335
DB 307 GKANATAMKONPETAKEIEKKVRELLLSNPSTP 341

RESULT 10
US-09-626-047-14
; Sequence 14, Application US/09626047
; Patent No. 6335198
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIA, VENKITSWARAN
; APPLICANT: CASTLE, LINDA
; APPLICANT: KREBER, CLAUD M.
; APPLICANT: BASS, STEVE
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE

; TITLE OF INVENTION: SEQUENCE RECOMBINATION
; FILE REFERENCE: 02-020720US
; CURRENT APPLICATION NUMBER: US/09/626,047
; CURRENT FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 09/116,188
; PRIOR FILING DATE: 07-15-1998
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: consensus
; OTHER INFORMATION: e. coli sequence
US-09-626-047-14

Query Match 100.0%; Score 1678; DB 2; Length 358;
Best Local Similarity 100.0%; Pred. No. 2.6e-167;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AIDENKQKALAAALGQIEKQFGKSGIMRLGEDRSMDVETISTGSLSLDIALGAGGLPMGR 60
DB 7 AIDENKQKALAAALGQIEKQFGKSGIMRLGEDRSMDVETISTGSLSLDIALGAGGLPMGR 66
QY 61 IVEIYGPESSGKTTTLTQVIAAAQREGKTCAPIDAEHALDPIYARKLGVDIDNLLCSQPD 120
DB 67 IVEIYGPESSGKTTTLTQVIAAAQREGKTCAPIDAEHALDPIYARKLGVDIDNLLCSQPD 126
QY 121 TGEQALEICDALARGAVDVIIVDSVAALTPKAEIEGEIGDISHMGLAARMMSQAMRKLKAG 180
DB 127 TGEQALEICDALARGAVDVIIVDSVAALTPKAEIEGEIGDISHMGLAARMMSQAMRKLKAG 186
QY 181 NLKQSNLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVG 240
DB 187 NLKQSNLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVG 246
QY 241 ETRVVKVKNKIAAPPKQAEFQILYGEINFGYELVDLGKELIEKAGAWYSYKGEKIQ 300
DB 247 ETRVVKVKNKIAAPPKQAEFQILYGEINFGYELVDLGKELIEKAGAWYSYKGEKIQ 306
QY 301 GKANATAMKONPETAKEIEKKVRELLLSNPSTP 335
DB 307 GKANATAMKONPETAKEIEKKVRELLLSNPSTP 341

RESULT 11
US-09-626-343-7
; Sequence 7, Application US/09626343
; Patent No. 6352859
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIA, VENKITSWARAN
; APPLICANT: CASTLE, LINDA
; APPLICANT: KREBER, CLAUD M.
; APPLICANT: BASS, STEVE
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
; FILE REFERENCE: 02-020720US
; CURRENT APPLICATION NUMBER: US/09/626,343
; CURRENT FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 09/116,188
; PRIOR FILING DATE: 07-15-1998
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 358
; TYPE: PRT

```
; ORGANISM: Escherichia coli
US-09-626-343-7

Query Match      100.0%; Score 1678; DB 2; Length 358;
Best Local Similarity 100.0%; Pred. No. 2.6e-167;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AIDENKOKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISGSLSLDIALGAGGLPMGR 60
DB 7 AIDENKOKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISGSLSLDIALGAGGLPMGR 66
QY 61 IVEIYGPESSGKTTLTQVIAAAQREGKTCAPDAEHALDPIYARKLGVDIDNLLCSQPD 120
DB 67 IVEIYGPESSGKTTLTQVIAAAQREGKTCAPDAEHALDPIYARKLGVDIDNLLCSQPD 126
QY 121 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKLKAG 180
DB 127 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKLKAG 186
QY 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVG 240
DB 187 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVG 246
QY 241 ETRVKVVKNKIAAPFKQAEFQILYGEINFGYELVDLGKVKELIEKAGAWISYKGEKIQ 300
DB 247 ETRVKVVKNKIAAPFKQAEFQILYGEINFGYELVDLGKVKELIEKAGAWISYKGEKIQ 306
QY 301 GKANATAWLKONPETAKEIEKKVRELLSNPNSTP 335
DB 307 GKANATAWLKONPETAKEIEKKVRELLSNPNSTP 341

RESULT 13
US-09-626-343-14
; Sequence 14, Application US/09626343
; Patent No. 6352859
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIA, VENKITSWARAN
; APPLICANT: CASTLE, LINDA
; APPLICANT: KREBBER, CLAUD M.
; APPLICANT: BASS, STEVE
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
; FILE REFERENCE: 02-020727US
; CURRENT APPLICATION NUMBER: US/09/626,343
; CURRENT FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 09/116,188
; PRIOR FILING DATE: 07-15-1998
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: consensus
; OTHER INFORMATION: e. coli sequence
US-09-626-343-14

Query Match      100.0%; Score 1678; DB 2; Length 358;
Best Local Similarity 100.0%; Pred. No. 2.6e-167;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AIDENKOKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISGSLSLDIALGAGGLPMGR 60
DB 7 AIDENKOKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISGSLSLDIALGAGGLPMGR 66
QY 61 IVEIYGPESSGKTTLTQVIAAAQREGKTCAPDAEHALDPIYARKLGVDIDNLLCSQPD 120
DB 67 IVEIYGPESSGKTTLTQVIAAAQREGKTCAPDAEHALDPIYARKLGVDIDNLLCSQPD 126
QY 121 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKLKAG 180
DB 127 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKLKAG 186
QY 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVG 240
DB 187 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVG 246
QY 241 ETRVKVVKNKIAAPFKQAEFQILYGEINFGYELVDLGKVKELIEKAGAWISYKGEKIQ 300
DB 247 ETRVKVVKNKIAAPFKQAEFQILYGEINFGYELVDLGKVKELIEKAGAWISYKGEKIQ 306

; ORGANISM: Escherichia coli
US-09-626-343-7

Query Match      100.0%; Score 1678; DB 2; Length 358;
Best Local Similarity 100.0%; Pred. No. 2.6e-167;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AIDENKOKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISGSLSLDIALGAGGLPMGR 60
DB 7 AIDENKOKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISGSLSLDIALGAGGLPMGR 66
QY 61 IVEIYGPESSGKTTLTQVIAAAQREGKTCAPDAEHALDPIYARKLGVDIDNLLCSQPD 120
DB 67 IVEIYGPESSGKTTLTQVIAAAQREGKTCAPDAEHALDPIYARKLGVDIDNLLCSQPD 126
QY 121 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKLKAG 180
DB 127 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKLKAG 186
QY 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVG 240
DB 187 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVG 246
QY 241 ETRVKVVKNKIAAPFKQAEFQILYGEINFGYELVDLGKVKELIEKAGAWISYKGEKIQ 300
DB 247 ETRVKVVKNKIAAPFKQAEFQILYGEINFGYELVDLGKVKELIEKAGAWISYKGEKIQ 306

; ORGANISM: Escherichia coli
US-09-626-343-7

Query Match      100.0%; Score 1678; DB 2; Length 358;
Best Local Similarity 100.0%; Pred. No. 2.6e-167;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AIDENKOKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISGSLSLDIALGAGGLPMGR 60
DB 7 AIDENKOKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISGSLSLDIALGAGGLPMGR 66
QY 61 IVEIYGPESSGKTTLTQVIAAAQREGKTCAPDAEHALDPIYARKLGVDIDNLLCSQPD 120
DB 67 IVEIYGPESSGKTTLTQVIAAAQREGKTCAPDAEHALDPIYARKLGVDIDNLLCSQPD 126
```

Db 247 ETRVVKVKNKIAAPFKQAEFOILYGEINFGYELVDLGVKEKLEKAGAWSYKGEKIGQ 306
Qy 301 GKANATAWKONPETAKEIEKKVRELLLSNPSTP 335
Db 307 GKANATAWKONPETAKEIEKKVRELLLSNPSTP 341

RESULT 14
US-09-354-922-8
; Sequence 8, Application US/09354922
; Patent No. 6379964
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIAN, VENKITESWATAN
; APPLICANT: CASTLE, LINDA A.
; APPLICANT: KREBER, CLAUD M.
; APPLICANT: BASS, STEVE
; APPLICANT: ZHANG, YING-XIN
; APPLICANT: COX, TONY
; APPLICANT: HUISMAN, GJALT
; APPLICANT: AFFHOLTER, JOSEPH A.
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
; FILE REFERENCE: 02-020730US
; CURRENT APPLICATION NUMBER: US/09/354,922
; CURRENT FILING DATE: 1999-07-15
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-354-922-8
Query Match 100.0%; Score 1678; DB 2; Length 358;
Best Local Similarity 100.0%; Pred. No. 2.6e-167;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AIDENKOKALAAALGOIEKQFGKSGIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 60
Db 7 AIDENKOKALAAALGOIEKQFGKSGIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 66
Qy 61 IVEIYGPESSGKTTTLQVIAAAQREGKTCAFIDAEHALDPIYARKLGVLDNLLCSQPD 120
Db 67 IVEIYGPESSGKTTTLQVIAAAQREGKTCAFIDAEHALDPIYARKLGVLDNLLCSQPD 126
Qy 121 TGEQALEICDALARSGAVDVIVDSVAALTPKAEIEGIGDSHMGLAARMMSQAMRKLKAG 180
Db 127 TGEQALEICDALARSGAVDVIVDSVAALTPKAEIEGIGDSHMGLAARMMSQAMRKLKAG 186
Qy 181 NLKOSNTLLIFINQIRMKIGVMFGNPETTTGCGNALKFYASVRLDIRRIGAVKEGENVVGS 240
Db 187 NLKOSNTLLIFINQIRMKIGVMFGNPETTTGCGNALKFYASVRLDIRRIGAVKEGENVVGS 246
Qy 241 ETRVVKVKNKIAAPFKQAEFOILYGEINFGYELVDLGVKEKLEKAGAWSYKGEKIGQ 300
Db 247 ETRVVKVKNKIAAPFKQAEFOILYGEINFGYELVDLGVKEKLEKAGAWSYKGEKIGQ 306
Qy 301 GKANATAWKONPETAKEIEKKVRELLLSNPSTP 335
Db 307 GKANATAWKONPETAKEIEKKVRELLLSNPSTP 341

RESULT 15
US-09-354-922-10
; Sequence 10, Application US/09354922
; Patent No. 6379964

; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIAN, VENKITESWATAN
; APPLICANT: CASTLE, LINDA A.
; APPLICANT: KREBER, CLAUD M.
; APPLICANT: BASS, STEVE
; APPLICANT: ZHANG, YING-XIN
; APPLICANT: COX, TONY
; APPLICANT: HUISMAN, GJALT
; APPLICANT: YUAN, LING
; APPLICANT: AFFHOLTER, JOSEPH A.
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
; FILE REFERENCE: 02-020730US
; CURRENT APPLICATION NUMBER: US/09/354,922
; CURRENT FILING DATE: 1999-07-15
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-354-922-10
Query Match 100.0%; Score 1678; DB 2; Length 358;
Best Local Similarity 100.0%; Pred. No. 2.6e-167;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AIDENKOKALAAALGOIEKQFGKSGIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 60
Db 7 AIDENKOKALAAALGOIEKQFGKSGIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 66
Qy 61 IVEIYGPESSGKTTTLQVIAAAQREGKTCAFIDAEHALDPIYARKLGVLDNLLCSQPD 120
Db 67 IVEIYGPESSGKTTTLQVIAAAQREGKTCAFIDAEHALDPIYARKLGVLDNLLCSQPD 126
Qy 121 TGEQALEICDALARSGAVDVIVDSVAALTPKAEIEGIGDSHMGLAARMMSQAMRKLKAG 180
Db 127 TGEQALEICDALARSGAVDVIVDSVAALTPKAEIEGIGDSHMGLAARMMSQAMRKLKAG 186
Qy 181 NLKOSNTLLIFINQIRMKIGVMFGNPETTTGCGNALKFYASVRLDIRRIGAVKEGENVVGS 240
Db 187 NLKOSNTLLIFINQIRMKIGVMFGNPETTTGCGNALKFYASVRLDIRRIGAVKEGENVVGS 246
Qy 241 ETRVVKVKNKIAAPFKQAEFOILYGEINFGYELVDLGVKEKLEKAGAWSYKGEKIGQ 300
Db 247 ETRVVKVKNKIAAPFKQAEFOILYGEINFGYELVDLGVKEKLEKAGAWSYKGEKIGQ 306
Qy 301 GKANATAWKONPETAKEIEKKVRELLLSNPSTP 335
Db 307 GKANATAWKONPETAKEIEKKVRELLLSNPSTP 341

Search completed: February 16, 2006, 01:26:08
Job time : 33.5 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 16, 2006, 01:39:50 ; Search time 123 Seconds
(without alignments)
1137.990 Million cell updates/sec

Title: US-10-733-782-1
Perfect score: 1678
Sequence: 1 AIDENKQKALAAALGOIEKQ.....AKEIEKKVRELLSNPNSTP 335

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA_Main:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1678	100.0	335	4 US-10-733-782-1	Sequence 1, Appli
2	1678	100.0	352	4 US-10-733-782-5	Sequence 5, Appli
3	1678	100.0	353	4 US-10-282-122A-43319	Sequence 43319, A
4	1678	100.0	358	4 US-10-194-686-8	Sequence 8, Appli
5	1678	100.0	358	4 US-10-194-686-10	Sequence 10, Appli
6	1678	100.0	358	4 US-10-194-686-15	Sequence 15, Appli
7	1674	99.8	335	4 US-10-733-782-3	Sequence 3, Appli
8	1673	99.7	358	4 US-10-194-686-9	Sequence 9, Appli
9	1673	99.7	358	4 US-10-194-686-12	Sequence 12, Appli
10	1660	99.9	358	4 US-10-194-686-13	Sequence 13, Appli
11	1659	98.9	358	4 US-10-194-686-11	Sequence 11, Appli
12	1641	97.8	353	4 US-10-282-122A-75948	Sequence 75948, A
13	1613	96.1	352	4 US-10-282-122A-59525	Sequence 59525, A
14	1593	94.9	352	4 US-10-282-122A-55632	Sequence 55632, A
15	1534	91.4	356	4 US-10-282-122A-78524	Sequence 78524, A
16	1504	89.6	355	4 US-10-282-122A-68593	Sequence 68593, A
17	1467	87.4	351	4 US-10-282-122A-72895	Sequence 72895, A
18	1436	85.6	412	4 US-10-282-122A-77075	Sequence 77075, A
19	1353	80.6	354	4 US-10-282-122A-67415	Sequence 67415, A
20	1332	79.4	354	4 US-10-282-122A-58235	Sequence 58235, A
21	1268	75.6	346	4 US-10-282-122A-66505	Sequence 66505, A
22	1259	75.0	349	4 US-10-282-122A-4768	Sequence 4768, A
23	1257	74.9	355	4 US-10-282-122A-67942	Sequence 67942, A
24	1255	74.8	369	4 US-10-282-122A-69802	Sequence 69802, A
25	1245.5	74.2	349	4 US-10-282-122A-63392	Sequence 63392, A
26	1241.5	74.0	348	4 US-10-282-122A-61049	Sequence 61049, A
27	1239.5	73.9	342	4 US-10-751-928-20	Sequence 20, Appli

ALIGNMENTS

RESULT 1

US-10-733-782-1
; Sequence 1, Application US/10733782
; Publication No. US20040157248A1
; GENERAL INFORMATION:
; APPLICANT: Cox, Michael
; APPLICANT: Lusetti, Shelley
; APPLICANT: Eggleter, Aimee
; TITLE OF INVENTION: RECA Mutants
; FILE REFERENCE: 960296.99501
; CURRENT APPLICATION NUMBER: US/10/733,782
; CURRENT FILING DATE: 2003-12-11
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 1
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-733-782-1

Query Match	100.0%;	Score 1678;	DB 4;	Length 335;
Best Local Similarity	100.0%;	Pred. No. 4.2e-152;		
Matches 335;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	AIDENKQKALAAALGOIEKQFGKGSIMRLGSDRSMVDVTIISTGSLSLDIALGAGGLPMGR	60	
Db	1	AIDENKQKALAAALGOIEKQFGKGSIMRLGSDRSMVDVTIISTGSLSLDIALGAGGLPMGR	60	
Qy	61	I VEIYGPSSCKTTLTLOVIAAQRGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD	120	
Db	61	I VEIYGPSSCKTTLTLOVIAAQRGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD	120	
Qy	121	TGEQALEICDALASGAVDVTVDSVAALTPKABIEGIGSHMGLAARMMSQAMRKLKAG	180	
Db	121	TGEQALEICDALASGAVDVTVDSVAALTPKABIEGIGSHMGLAARMMSQAMRKLKAG	180	
Qy	181	NLKQSNLTLLIFINQIRMKIGVMFQNPETTGTGNALKFYASVRLDIRRIGAVKEGNNVVS	240	
Db	181	NLKQSNLTLLIFINQIRMKIGVMFQNPETTGTGNALKFYASVRLDIRRIGAVKEGNNVVS	240	
Qy	241	ETRVKVVKNKTAAPFKQAEFQILYGEFINFGELVDLGKLEKAGAWYSYKGEKIQG	300	
Db	241	ETRVKVVKNKTAAPFKQAEFQILYGEFINFGELVDLGKLEKAGAWYSYKGEKIQG	300	
Qy	301	GKANATAWLKONPETAKEIEKKVRELLSNPNSTP	335	
Db	301	GKANATAWLKONPETAKEIEKKVRELLSNPNSTP	335	

RESULT 2

```
US-10-733-782-5
; Sequence 5, Application US/10733782
; Publication No. US20040157248A1
; GENERAL INFORMATION:
; APPLICANT: Cox, Michael
; APPLICANT: Lusetti, Shelley
; APPLICANT: Egger, Almee
; TITLE OF INVENTION: RecA Mutants
; FILE REFERENCE: 960296.99501
; CURRENT APPLICATION NUMBER: US/10/733,782
; CURRENT FILING DATE: 2003-12-11
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-733-782-5
Query Match 100.0%; Score 1678; DB 4; Length 352;
Best Local Similarity 100.0%; Pred. No. 4.5e-152;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AIDENKQKALAAALQIEKQFGKSGIMRLGEDRSMVDVETISTGSLSDIALGAGGLPMGR 60
Db 1 AIDENKQKALAAALQIEKQFGKSGIMRLGEDRSMVDVETISTGSLSDIALGAGGLPMGR 60
QY 61 IVEIYGPESGKTTTLQVIAAAQREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 61 IVEIYGPESGKTTTLQVIAAAQREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
QY 121 TGEQALEICDALARSGAVDVIVVDSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKLKAG 180
Db 121 TGEQALEICDALARSGAVDVIVVDSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKLKAG 180
QY 181 NLKQSTLLIFINQIRMKIGWFGNPNPTTGGNALKFYASVRLDIRRIGAVGENNVGS 240
Db 181 NLKQSTLLIFINQIRMKIGWFGNPNPTTGGNALKFYASVRLDIRRIGAVGENNVGS 240
QY 241 ETRVKVVKNKIAAPFKQAEFQILYGEINFGYBLVDLGVKEKLIKAGAWYSYKGEKIGQ 300
Db 241 ETRVKVVKNKIAAPFKQAEFQILYGEINFGYBLVDLGVKEKLIKAGAWYSYKGEKIGQ 300
QY 301 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 335
Db 301 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 335
RESULT 3
US-10-733-782-5
; Sequence 43319, Application US/10282122A
; Publication No. US2004029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zvekind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
US-10-733-782-1
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 43319
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-282-122A-43319
Query Match 100.0%; Score 1678; DB 4; Length 353;
Best Local Similarity 100.0%; Pred. No. 4.5e-152;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AIDENKQKALAAALQIEKQFGKSGIMRLGEDRSMVDVETISTGSLSDIALGAGGLPMGR 60
Db 2 AIDENKQKALAAALQIEKQFGKSGIMRLGEDRSMVDVETISTGSLSDIALGAGGLPMGR 61
QY 61 IVEIYGPESGKTTTLQVIAAAQREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 62 IVEIYGPESGKTTTLQVIAAAQREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 121
QY 121 TGEQALEICDALARSGAVDVIVVDSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKLKAG 180
Db 122 TGEQALEICDALARSGAVDVIVVDSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKLKAG 181
QY 181 NLKQSTLLIFINQIRMKIGWFGNPNPTTGGNALKFYASVRLDIRRIGAVGENNVGS 240
Db 182 NLKQSTLLIFINQIRMKIGWFGNPNPTTGGNALKFYASVRLDIRRIGAVGENNVGS 241
QY 241 ETRVKVVKNKIAAPFKQAEFQILYGEINFGYBLVDLGVKEKLIKAGAWYSYKGEKIGQ 300
Db 242 ETRVKVVKNKIAAPFKQAEFQILYGEINFGYBLVDLGVKEKLIKAGAWYSYKGEKIGQ 301
QY 301 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 335
Db 302 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 336
RESULT 4
US-10-194-686-8
; Sequence 8, Application US/10194686
; Publication No. US20030148309A1
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIAN, VENKITESWATAN
; APPLICANT: CASTLE, LINDA A.
; APPLICANT: KREBER, CLAUS M.
; APPLICANT: BASS, STEVE
; APPLICANT: ZHANG, YING-XIN
; APPLICANT: COX, TONY
; APPLICANT: HUISMAN, GJALT
; APPLICANT: YUAN, LING
```

```
; APPLICANT: AFFHOLTER, JOSEPH A.
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
; FILE REFERENCE: 02-020730US
; CURRENT APPLICATION NUMBER: US/10/194,686
; PRIOR FILING DATE: 2002-07-11
; PRIOR APPLICATION NUMBER: US/09/354,922
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-194-686-8

Query Match          100.0%; Score 1678; DB 4; Length 358;
Best Local Similarity 100.0%; Pred. No. 4.6e-152; Indels 0; Gaps 0;
Matches 335; Conservative 0; Mismatches 0;

Qy 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGDRSDMDVETISTGSLSLDIALGAGGLPMGR 60
Db 7 AIDENKQKALAAALGQIEKQFGKGSIMRLGDRSDMDVETISTGSLSLDIALGAGGLPMGR 66

Qy 61 IVEIYGPESSGKTTTLQVIAAQAQREGKTCFIDAEHALDPIYARKLGVLDNLLCSQPD 120
Db 67 IVEIYGPESSGKTTTLQVIAAQAQREGKTCFIDAEHALDPIYARKLGVLDNLLCSQPD 126

Qy 121 TGEQALEICDALARSGADVIVVDSVAALTPKAEIEGIEGDSHMGSLAARMMSQAMRKLKAG 180
Db 127 TGEQALEICDALARSGADVIVVDSVAALTPKAEIEGIEGDSHMGSLAARMMSQAMRKLKAG 186

Qy 181 NLKQSNLTLLIIFINQIRMKIGVMFGNPETTTGKNALKFYASVRLDIRRIGAVKEGNNVVG 240
Db 187 NLKQSNLTLLIIFINQIRMKIGVMFGNPETTTGKNALKFYASVRLDIRRIGAVKEGNNVVG 246

Qy 241 ETRVKVKNKIAAPFKQAEFOILYGEINFGYELVDLGKVKLEKAGAWYSYKGEKIGQ 300
Db 247 ETRVKVKNKIAAPFKQAEFOILYGEINFGYELVDLGKVKLEKAGAWYSYKGEKIGQ 306

Qy 301 GKANATWLKONPETAKEIEKKVRELLLSNPSTP 335
Db 307 GKANATWLKONPETAKEIEKKVRELLLSNPSTP 341

RESULT 5
US-10-194-686-10
; Sequence 10, Application US/10194686
; Publication No. US20030148309A1
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIAN, VENKITESWATAN
; APPLICANT: CASTLE, LINDA A.
; APPLICANT: KREBER, CLAUD M.
; APPLICANT: BASS, STEVE
; APPLICANT: ZHANG, YING-XIN
; APPLICANT: COX, TONY
; APPLICANT: HUISMAN, GJALT
; APPLICANT: YUAN, LING
; APPLICANT: AFFHOLTER, JOSEPH A.
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
; FILE REFERENCE: 02-020730US
; CURRENT APPLICATION NUMBER: US/10/194,686
; PRIOR FILING DATE: 2002-07-11
; PRIOR APPLICATION NUMBER: US/09/354,922
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: consensus
; OTHER INFORMATION: e. coli sequence
US-10-194-686-15
```

```
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-194-686-10

Query Match          100.0%; Score 1678; DB 4; Length 358;
Best Local Similarity 100.0%; Pred. No. 4.6e-152; Indels 0; Gaps 0;
Matches 335; Conservative 0; Mismatches 0;

Qy 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGDRSDMDVETISTGSLSLDIALGAGGLPMGR 60
Db 7 AIDENKQKALAAALGQIEKQFGKGSIMRLGDRSDMDVETISTGSLSLDIALGAGGLPMGR 66

Qy 61 IVEIYGPESSGKTTTLQVIAAQAQREGKTCFIDAEHALDPIYARKLGVLDNLLCSQPD 120
Db 67 IVEIYGPESSGKTTTLQVIAAQAQREGKTCFIDAEHALDPIYARKLGVLDNLLCSQPD 126

Qy 121 TGEQALEICDALARSGADVIVVDSVAALTPKAEIEGIEGDSHMGSLAARMMSQAMRKLKAG 180
Db 127 TGEQALEICDALARSGADVIVVDSVAALTPKAEIEGIEGDSHMGSLAARMMSQAMRKLKAG 186

Qy 181 NLKQSNLTLLIIFINQIRMKIGVMFGNPETTTGKNALKFYASVRLDIRRIGAVKEGNNVVG 240
Db 187 NLKQSNLTLLIIFINQIRMKIGVMFGNPETTTGKNALKFYASVRLDIRRIGAVKEGNNVVG 246

Qy 241 ETRVKVKNKIAAPFKQAEFOILYGEINFGYELVDLGKVKLEKAGAWYSYKGEKIGQ 300
Db 247 ETRVKVKNKIAAPFKQAEFOILYGEINFGYELVDLGKVKLEKAGAWYSYKGEKIGQ 306

Qy 301 GKANATWLKONPETAKEIEKKVRELLLSNPSTP 335
Db 307 GKANATWLKONPETAKEIEKKVRELLLSNPSTP 341

RESULT 6
US-10-194-686-15
; Sequence 15, Application US/10194686
; Publication No. US20030148309A1
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIAN, VENKITESWATAN
; APPLICANT: CASTLE, LINDA A.
; APPLICANT: KREBER, CLAUD M.
; APPLICANT: BASS, STEVE
; APPLICANT: ZHANG, YING-XIN
; APPLICANT: COX, TONY
; APPLICANT: HUISMAN, GJALT
; APPLICANT: YUAN, LING
; APPLICANT: AFFHOLTER, JOSEPH A.
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
; FILE REFERENCE: 02-020730US
; CURRENT APPLICATION NUMBER: US/10/194,686
; PRIOR FILING DATE: 2002-07-11
; PRIOR APPLICATION NUMBER: US/09/354,922
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: consensus
; OTHER INFORMATION: e. coli sequence
US-10-194-686-15
```

Query Match 100.0%; Score 1678; DB 4; Length 358;
Best Local Similarity 100.0%; Pred. No. 4.6e-152;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 60
DB 7 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 66

QY 61 IVEIYGPSSGKTTTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
DB 67 IVEIYGPSSGKTTTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 126

QY 121 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGEIGDSDHMGGLAARMMSQAMRKLKAG 180
DB 127 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGEIGDSDHMGGLAARMMSQAMRKLKAG 186

QY 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENNVGS 240
DB 187 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENNVGS 246

QY 241 ETRVKVVKNKIAAPFKQAEFQILYGEINFGYELVDLGKVKELIEKAGAWYSYKGEKIGQ 300
DB 247 ETRVKVVKNKIAAPFKQAEFQILYGEINFGYELVDLGKVKELIEKAGAWYSYKGEKIGQ 306

QY 301 GKANATAWLKNDPETAKEIEKKVRELLLSNPSTP 335
DB 307 GKANATAWLKNDPETAKEIEKKVRELLLSNPSTP 341

RESULT 7
US-10-733-782-3
; Sequence 3, Application US/10733782
; Publication No. US20040157248A1
; GENERAL INFORMATION:
; APPLICANT: Cox, Michael
; APPLICANT: Lusetti, Shelley
; APPLICANT: Egger, Aimee
; TITLE OF INVENTION: RecA Mutants
; FILE REFERENCE: 960296.99501
; CURRENT APPLICATION NUMBER: US/10/733,782
; CURRENT FILING DATE: 2003-12-11
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-733-782-3

Query Match 99.8%; Score 1674; DB 4; Length 335;
Best Local Similarity 99.7%; Pred. No. 1e-151;
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 60
DB 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 60

QY 61 IVEIYGPSSGKTTTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
DB 61 IVEIYGPSSGKTTTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 120

QY 121 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGEIGDSDHMGGLAARMMSQAMRKLKAG 180
DB 121 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGEIGDSDHMGGLAARMMSQAMRKLKAG 180

QY 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENNVGS 240
DB 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENNVGS 240

QY 241 ETRVKVVKNKIAAPFKQAEFQILYGEINFGYELVDLGKVKELIEKAGAWYSYKGEKIGQ 300
DB 241 ETRVKVVKNKIAAPFKQAEFQILYGEINFGYELVDLGKVKELIEKAGAWYSYKGEKIGQ 300

QY 301 GKANATAWLKNDPETAKEIEKKVRELLLSNPSTP 335
DB 301 GKANATAWLKNDPETAKEIEKKVRELLLSNPSTP 335

RESULT 8
US-10-194-686-9
; Sequence 9, Application US/10194686
; Publication No. US20030148309A1
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P. C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIAN, VENKITESWATAN
; APPLICANT: CASTLE, LINDA A.
; APPLICANT: KREBER, CLAU M.
; APPLICANT: BASS, STEVE
; APPLICANT: ZHANG, YING-XIN
; APPLICANT: COX, TONY
; APPLICANT: HUISMAN, GJALT
; APPLICANT: YUAN, LING
; APPLICANT: AFFHOLTER, JOSEPH A.
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
; FILE REFERENCE: 02-020730US
; CURRENT APPLICATION NUMBER: US/10/194,686
; CURRENT FILING DATE: 2002-07-11
; PRIOR APPLICATION NUMBER: US/09/354,922
; PRIOR FILING DATE: 1999-07-15
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-194-686-9

Query Match 99.7%; Score 1673; DB 4; Length 358;
Best Local Similarity 99.4%; Pred. No. 1.4e-151;
Matches 333; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 60
DB 7 AIDENKQKALATAGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 66

QY 61 IVEIYGPSSGKTTTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
DB 67 IVEIYGPSSGKTTTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 126

QY 121 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGEIGDSDHMGGLAARMMSQAMRKLKAG 180
DB 127 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGEIGDSDHMGGLAARMMSQAMRKLKAG 186

QY 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENNVGS 240
DB 187 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENNVGS 246

QY 241 ETRVKVVKNKIAAPFKQAEFQILYGEINFGYELVDLGKVKELIEKAGAWYSYKGEKIGQ 300
DB 247 ETRVKVVKNKIAAPFKQAEFQILYGEINFGYELVDLGKVKELIEKAGAWYSYKGEKIGQ 306

QY 301 GKANATAWLKNDPETAKEIEKKVRELLLSNPSTP 335
DB 307 GKANATAWLKNDPETAKEIEKKVRELLLSNPSTP 341

RESULT 9
US-10-194-686-12
; Sequence 12, Application US/10194686

```
; Publication No. US20030148309A1
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIAN, VENKITESWATAN
; APPLICANT: CASTLE, LINDA A.
; APPLICANT: KREBER, CLAUD M.
; APPLICANT: BASS, STEVE
; APPLICANT: ZHANG, YING-XIN
; APPLICANT: AFFHOLTER, JOSEPH A.
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
; FILE REFERENCE: 02-020730US
; CURRENT APPLICATION NUMBER: US/10/194,686
; PRIOR FILING DATE: 2002-07-11
; PRIOR APPLICATION NUMBER: US/09/354,922
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Escherichia coli
; US-10-194-686-12

Query Match          99.7%; Score 1673; DB 4; Length 358;
Best Local Similarity 99.4%; Pred. No. 1.4e-151;
Matches 333; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 AIDENKQKALAAALGOIEKQFGKSGIMRLGEDRSDMDVETITSGSLSLDIALGAGGLPMGR 60
DB 7 AIDENKQKALAAALGOIEKQFGKSGIMRLGEDRSDMDVETITSGSLSLDIALGAGGLPMGR 66

QY 61 IVEIYGPESSGKTTTLQVIAAAQREGKTCAPIDAEHALDPIYARKLGVLDINLLCSQPD 120
DB 67 IVEIYGPESSGKTTTLQVIAAAQREGKTCAPIDAEHALDPIYARKLGVLDINLLCSQPD 126

QY 121 TGEQALEICDALSARGVDVIVDSVAALTPKAEIEGEGISHMGLAARMWSQAMRKLKAG 180
DB 127 TGEQALEICDALSARGVDVIVDSVAALTPKAEIEGEGISHMGLAARMWSQAMRKLKAG 186

QY 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGNNALKFYASVRLDIRRIGAVKEGENVVGS 240
DB 187 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGNNALKFYASVRLDIRRIGAVKEGENVVGS 246

QY 241 ETRVKVKNKIAAPFKQAEFOILYGEINFGYELVDLGKVKLEIKAGAWTSYKGEKIGQ 300
DB 247 ETRVKVKNKIAAPFKQAEFOILYGEINFGYELVDLGKVKLEIKAGAWTSYKGEKIGQ 306

QY 301 GKANATWLKONPETAKEIEKKVRELLSNPNSTP 335
DB 307 GKANATWLKONPETAKEIEKKVRELLSNPNSTP 341

RESULT 10
US-10-194-686-13
; Sequence 13, Application US/10194686
; Publication No. US20030148309A1
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIAN, VENKITESWATAN
```

```
; APPLICANT: CASTLE, LINDA A.
; APPLICANT: KREBER, CLAUD M.
; APPLICANT: BASS, STEVE
; APPLICANT: ZHANG, YING-XIN
; APPLICANT: COX, TONY
; APPLICANT: HUISMAN, GJALT
; APPLICANT: YUAN, LING
; APPLICANT: AFFHOLTER, JOSEPH A.
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
; FILE REFERENCE: 02-020730US
; CURRENT APPLICATION NUMBER: US/10/194,686
; PRIOR FILING DATE: 2002-07-11
; PRIOR APPLICATION NUMBER: US/09/354,922
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Escherichia coli
; US-10-194-686-13

Query Match          98.9%; Score 1660; DB 4; Length 358;
Best Local Similarity 98.8%; Pred. No. 2.5e-150;
Matches 331; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 AIDENKQKALAAALGOIEKQFGKSGIMRLGEDRSDMDVETITSGSLSLDIALGAGGLPMGR 60
DB 7 AIDENKQKALAAALGOIEKQFGKSGIMRLGEDRSDMDVETITSGSLSLDIALGAGGLPMGR 66

QY 61 IVEIYGPESSGKTTTLQVIAAAQREGKTCAPIDAEHALDPIYARKLGVLDINLLCSQPD 120
DB 67 IVEIYGPESSGKTTTLQVIAAAQREGKTCAPIDAEHALDPIYARKLGVLDINLLCSQPD 126

QY 121 TGEQALEICDALSARGVDVIVDSVAALTPKAEIEGEGISHMGLAARMWSQAMRKLKAG 180
DB 127 TGEQALEICDALSARGVDVIVDSVAALTPKAEIEGEGISHMGLAARMWSQAMRKLKAG 186

QY 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGNNALKFYASVRLDIRRIGAVKEGENVVGS 240
DB 187 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGNNALKFYASVRLDIRRIGAVKEGENVVGS 246

QY 241 ETRVKVKNKIAAPFKQAEFOILYGEINFGYELVDLGKVKLEIKAGAWTSYKGEKIGQ 300
DB 247 ETRVKVKNKIAAPFKQAEFOILYGEINFGYELVDLGKVKLEIKAGAWTSYKGEKIGQ 306

QY 301 GKANATWLKONPETAKEIEKKVRELLSNPNSTP 335
DB 307 GKANATWLKONPETAKEIEKKVRELLSNPNSTP 341

RESULT 11
US-10-194-686-11
; Sequence 11, Application US/10194686
; Publication No. US20030148309A1
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIAN, VENKITESWATAN
; APPLICANT: CASTLE, LINDA A.
; APPLICANT: KREBER, CLAUD M.
; APPLICANT: BASS, STEVE
; APPLICANT: ZHANG, YING-XIN
; APPLICANT: COX, TONY
; APPLICANT: HUISMAN, GJALT
; APPLICANT: YUAN, LING
; APPLICANT: AFFHOLTER, JOSEPH A.
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
```

Thu Feb 16 11:15:02 2006

```

; TITLE OF INVENTION: SEQUENCE RECOMBINATION
; FILE REFERENCE: 02-020730US
; CURRENT APPLICATION NUMBER: US/10/194,686
; CURRENT FILING DATE: 2002-07-11
; PRIOR APPLICATION NUMBER: US/09/354,922
; PRIOR FILING DATE: 1999-07-15
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-194-686-11

Query Match      98.9%; Score 1659; DB 4; Length 358;
Best Local Similarity 99.1%; Pred. No. 3.1e-150;
Matches 332; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 60
DB 7 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 66
QY 61 IVEIYGPSSSGKTTTLQVIAAAQREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
DB 67 IVEIYGPSSSGKTTTLQVIAAAQREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 126
QY 121 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKLAG 180
DB 127 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKLAG 186
QY 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVGS 240
DB 187 NLKLSNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVGS 246
QY 241 ETRVKVVKXKIAAPPKQAEFQILYGEINFGYELVDLGVKEKLEKAGAWYSYKGEKIQ 300
DB 247 ETRVKVVKXKIAAPPKQAEFQILYGEINFGYELVDLGVKEKLEKAGAWYSYKGEKIQ 306
QY 301 GKANATAWLKONPETAKEIEKKVRELLSNPNSTP 335
DB 307 GKANAALWLNKONPETAKEIEKKVRELLSNPNSTP 341

RESULT 12
US-10-282-122A-75948
; Sequence 75948, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347

; TITLE OF INVENTION: SEQUENCE RECOMBINATION
; FILE REFERENCE: 02-020730US
; CURRENT APPLICATION NUMBER: US/10/194,686
; CURRENT FILING DATE: 2002-07-11
; PRIOR APPLICATION NUMBER: US/09/354,922
; PRIOR FILING DATE: 1999-07-15
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 75948
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Salmonella typhi
US-10-282-122A-75948

Query Match      97.8%; Score 1641; DB 4; Length 353;
Best Local Similarity 97.3%; Pred. No. 1.6e-148;
Matches 326; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 60
DB 2 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVETISTGSLSLDIALGAGGLPMGR 61
QY 61 IVEIYGPSSSGKTTTLQVIAAAQREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
DB 62 IVEIYGPSSSGKTTTLQVIAAAQREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 121
QY 121 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKLAG 180
DB 122 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKLAG 181
QY 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVGS 240
DB 182 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVGS 241
QY 241 ETRVKVVKXKIAAPPKQAEFQILYGEINFGYELVDLGVKEKLEKAGAWYSYKGEKIQ 300
DB 242 ETRVKVVKXKIAAPPKQAEFQILYGEINFGYELVDLGVKEKLEKAGAWYSYKGEKIQ 301
QY 301 GKANATAWLKONPETAKEIEKKVRELLSNPNSTP 335
DB 302 GKANATTLWLNKONPETAKEIEKKVRELLSNPNSTP 336

RESULT 13
US-10-282-122A-59525
; Sequence 59525, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
```



```

; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 59525
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-10-282-122A-59525

Query Match
Best Local Similarity 96.1%; Score 1613; DB 4; Length 352;
Matches 320; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

QY 1 AIDENKQKALAALQIEKQFGKGSIMRLGEDRSDMDVETISTGSLSLDIALGAGGLPMGR 60
Db 2 AIDENKQKALAALQIEKQFGKGSIMRLGEDRSDMDVETISTGSLSLDIALGAGGLPMGR 61
QY 61 IVEIYGPSSGKTTLTQVIAAAQREGKTCAPIDAEHALDPYARKLGVDIDNLLCSQPD 120
Db 62 IVEIYGPSSGKTTLTQVIAAAQREGKTCAPIDAEHALDPYARKLGVDIDNLLCSQPD 121
QY 121 TGEQALEICDALARSGVDVIVDSVAALTPKAEIEGEIGDGHMGLAARMMSQAMRKLK 180
Db 122 TGEQALEICDALARSGVDVIVDSVAALTPKAEIEGEIGDGHMGLAARMMSQAMRKLK 181
QY 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGNNALKFYASVRLDIRRIGAVKEGNNVGS 240
Db 182 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGNNALKFYASVRLDIRRIGAVKEGNNVGS 241
QY 241 ETRVKVKNKTAAPFKQAEFOILYGEINFGYELVDLGVKELIEKAGAWTSYKGEKIQ 300
Db 242 ETRVKVKNKTAAPFKQAEFOILYGEINFGYELVDLGVKELIEKAGAWTSYKGEKIQ 301
QY 301 GKANATWLKONPETAKEIEKKVRELLLSNPSTP 335
Db 302 GKANAITWLKENPAAAKEIEKKVRELLLNQDQAKP 336

RESULT 14
US-10-282-122A-55632
; Sequence 55632, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 55632
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Enterobacter cloacae
US-10-282-122A-55632

Query Match
Best Local Similarity 94.9%; Score 1593; DB 4; Length 352;
Matches 317; Conservative 9; Mismatches 9; Indels 0; Gaps 0;

QY 1 AIDENKQKALAALQIEKQFGKGSIMRLGEDRSDMDVETISTGSLSLDIALGAGGLPMGR 60
Db 2 AIDENKQKALAALQIEKQFGKGSIMRLGEDRSDMDVETISTGSLSLDIALGAGGLPMGR 61
QY 61 IVEIYGPSSGKTTLTQVIAAAQREGKTCAPIDAEHALDPYARKLGVDIDNLLCSQPD 120
Db 62 IVEIYGPSSGKTTLTQVIAAAQREGKTCAPIDAEHALDPYARKLGVDIDNLLCSQPD 121
QY 121 TGEQALEICDALARSGVDVIVDSVAALTPKAEIEGEIGDGHMGLAARMMSQAMRKLK 180
Db 122 TGEQALEICDALARSGVDVIVDSVAALTPKAEIEGEIGDGHMGLAARMMSQAMRKLK 181
QY 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGNNALKFYASVRLDIRRIGAVKEGNNVGS 240
Db 182 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGNNALKFYASVRLDIRRIGAVKEGNNVGS 241
QY 241 ETRVKVKNKTAAPFKQAEFOILYGEINFGYELVDLGVKELIEKAGAWTSYKGEKIQ 300
Db 242 ETRVKVKNKTAAPFKQAEFOILYGEINFGYELVDLGVKELIEKAGAWTSYKGEKIQ 301
QY 301 GKANATWLKONPETAKEIEKKVRELLLSNPSTP 335
Db 302 GKANAISWLKENPAAAKEIEKKVRELLLNQDQSKP 336

RESULT 15
US-10-282-122A-78524
; Sequence 78524, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
```

Thu Feb 16 11:15:02 2006

```

; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 78524
; LENGTH: 356
; TYPE: PRT
; ORGANISM: Yersinia pestis
; US-10-282-122A-78524

Query Match          91.4%; Score 1534; DB 4; Length 356;
Best Local Similarity 91.9%; Pred. No. 3e-138;
Matches 305; Conservative 14; Mismatches 13; Indels 0; Gaps 0;

QY 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMDEVETISTGSLSDIALGAGGLPMGR 60
Db 2 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMDEVETISTGSLSDIALGAGGLPMGR 61

QY 61 IVEIYGPSSGKTTLTQVIAAAREGKTCAFDAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 62 IVEIYGPSSGKTTLTQVIAAAREGKTCAFDAEHALDPIYARKLGVDIDNLLCSQPD 121

QY 121 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKL 180
Db 122 TGEQALEICDALTFSGNVDVIIIVDSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKL 181

QY 181 NLKQSNLTLLIFINQIRMKIGWFGNPTTTGGNALKFYASVRLDIRRIGAVKEGENVVG 240
Db 182 NLKANTLLIFINQIRMKIGWFGNPTTTGGNALKFYASVRLDIRRIGAVKDGDDVVVG 241

QY 241 ETRVKVYKNIAPFFQAEQILYGEINFGYELVDLGVEKLEKAGAWYSYKGEKIG 300
Db 242 ETRVKVYKNIAPFFQAEQILYGEINFGYELVDLGVEKLEKAGAWYSYKGEKIG 301

QY 301 GKANATAWKNDPETAKEIEKKVRELLLSNPN 332
Db 302 GKANASNYLKENPAIAAELOKUREMLLNGN 333
```

Search completed: February 16, 2006, 01:44:34
Job time : 124 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 16, 2006, 01:40:35 ; Search time 10.5 Seconds
(without alignments)
453.421 Million cell updates/sec

Title: US-10-733-782-1
Perfect score: 1678
Sequence: 1 AIDENKQKALAAALQIEKQ.....AKEIEKKVRELLLSNPNTSP 335

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 107799 seqs, 14211699 residues

Total number of hits satisfying chosen parameters: 107799

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA_New.*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US05_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1212.5	72.3	375	6	US-10-467-657-4766	Sequence 4766, Ap
2	1019	60.7	363	7	US-11-074-176-296	Sequence 296, App
3	970.5	57.8	355	7	US-11-098-686-11427	Sequence 11427, A
4	155	9.2	443	7	US-11-098-686-10861	Sequence 10861, A
5	111.5	6.6	459	6	US-10-467-657-3092	Sequence 3092, App
6	109	6.5	453	6	US-10-467-657-206	Sequence 206, App
7	109	6.5	453	6	US-10-467-657-6400	Sequence 6400, Ap
8	104	6.2	454	7	US-11-098-686-10599	Sequence 10599, A
9	102	6.1	453	6	US-10-467-657-3626	Sequence 3626, Ap
10	101	6.0	425	6	US-10-454-437-50	Sequence 50, Appl
11	100	6.0	471	6	US-10-467-657-802	Sequence 802, App
12	99	5.9	421	6	US-10-467-657-4148	Sequence 4148, Ap
13	97	5.8	365	7	US-11-082-389-132	Sequence 132, App
14	97	5.8	376	7	US-11-213-368-14	Sequence 14, Appl
15	97	5.8	479	7	US-11-082-389-130	Sequence 130, App
16	96	5.7	286	7	US-11-082-389-88	Sequence 88, Appl
17	95.5	5.7	419	6	US-10-979-821-6	Sequence 6, Appli
18	95.5	5.7	419	7	US-11-114-922-6	Sequence 6, Appli
19	95	5.7	257	7	US-11-156-084-264	Sequence 264, App
20	95	5.7	376	7	US-11-082-389-122	Sequence 122, App
21	95	5.7	376	7	US-11-082-389-124	Sequence 124, App
22	94.5	5.6	640	7	US-11-205-109-9	Sequence 9, Appli
23	94.5	5.6	716	6	US-10-131-826A-96	Sequence 96, Appl
24	94	5.6	406	6	US-10-467-657-7420	Sequence 7420, Ap
25	94	5.6	655	7	US-11-094-586-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1

US-10-467-657-4766
; Sequence 4766, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON Spa
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 4766
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
; US-10-467-657-4766

Query Match	72.3%	Score 1212.5;	DB 6;	Length 375;
Best Local Similarity	69.0%	Pred. No. 4e-86;		
Matches	225;	Conservative 52;	Mismatches 48;	Indels 1; Gaps 1;
Qy	2	IDENKQKALAAALQIEKQFGKGSIMRL-GEGRSDMDVETISTGSLSLDIALGAGGLPMGR	60	
Db	28	MSDDSKALAAALQIEKSGFKGAIKMGDSQOELEVIETGSLGLDLALGVGGLPRGR	87	
Qy	61	IIVELGPSSCKTTLTQVIAAQAQREGKTCAFIDAEHALDPIYARKLGVDNLLCSQPD	120	
Db	88	IIVELGPSSCKTTLTCLAEVAQAQCKNGGVCAFPVDAEHAFDPVYARKLGKVEEYLSQPD	147	
Qy	121	TGEQALEICDALARSGAVDVIWVDSVAALTPKATIEGIGSHMGLAARMMSQAMRKLKAG	180	
Db	148	TGEQALEICDVLVSSGGIDVWVDSVAALTVPKATIEGDMGSHVGLQARLMSQALRKLKG	207	
Qy	181	NLKQSTLLIPINQIRMKIGVMFGNPETTTGNCALKFYASVRLDIRRTGAVKEGNNVGS	240	
Db	208	HIKKTNTLVVFINQIRMKIGVMFGSPETTTGNCALKFYSSVRLDIRRTGSIKKGSEVLGN	267	
Qy	241	ETRVKVNKNTAAPFKQAEFOILYCEGINFYGELVDLGKLEKTEKAGAWYSYKEKIGQ	300	
Db	268	ETRVKVNKNVAPPFRQAEFDILYEGISWEGELIDIGVKNDIINKSGAWYSYNGAKITG	327	

QY 136 GAVDVIVDVSAALTPKABIEGEGSHMGLAARMQSMQMRKLAGNL-----KQSNLTLLIF 191
Db 160 EAPDLLLDSVQTLT--SSNAELPGN-----VSQ-VRAVATELVBEACQKGTTVIF 208
QY 192 INQIRMKIGVMFG 204
Db 209 IGHV-TKDGTLAG 220

RESULT 5

US-10-467-657-3092
; Sequence 3092, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON Spa
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 3092
; LENGTH: 459
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-3092

Query Match 6.6%; Score 111.5; DB 6; Length 459;

Best Local Similarity 22.0%; Pred. No. 0.23;
Matches 65; Conservative 44; Mismatches 127; Indels 59; Gaps 12;
QY 25 SIMRIGEDRSMVDVETISTGSLSLDIALGAGGLPMGRIVEIYGPSSGKTTTLQVIAAAQ 84
Db 57 TVQSLSAVTAABVPRNPTGMGELDRVLG--GGLVDGAVILLGGDPGIGKSTLLQITAKWA 115
QY 85 REGKTCATDAEHALDPYARKLVGDID----NLLCSQPDGTGEQALEICDALARSGVDV 140
Db 116 QSRKV-LVVSSESQAQVALRSQRLEHAEGVNL-----ABIRMEAIQRAALKQHQPEV 168
QY 141 IWVDSVAALTPKAEIEGEGSHMGLAARMQSMQMRKLAGNL-----KQSNLTLLIFINQIR 196
Db 169 VVIDSIQWY-----SDQITSAPGSVSQ-VRECAQLTRWAKQWGIAMILLVGHV- 216
QY 197 MKIGVMFGNPETT-----GGNALKFYASVRLDIRRIGAVKEGENVVVGSETRVKVW 247
Db 217 TKDGAAG-PRVLEHMDVTLVYFEGDQHSNYRMIRAIKNRFGAANELGVFAWTEGLKGV 275
QY 248 KAKIA---APFKQ-----ABFOILYGEGINFYGELVDLGKZE 281
Db 276 SNPSAIFLASYRDDTFPGSCVLVTQSGRPLLVIEIQALVDDAHGFTPKRLSVGLEQ 330

RESULT 6

US-10-467-657-206
; Sequence 206, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON Spa
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8

; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 206
; LENGTH: 453
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-206

Query Match 6.5%; Score 109; DB 6; Length 453;

Best Local Similarity 23.5%; Pred. No. 0.36;
Matches 65; Conservative 43; Mismatches 91; Indels 78; Gaps 16;
QY 1 AIDENKOKALAAALGQ-IEKQFGKGSIMRLGEDRSMVDVET--ISTGSLSLDIALG-AGGL 56
Db 151 AVGDVAVKRETKTQGTVEDLIG-----GLDKRLDGVRFGLPTGLMKLD---GWTGGL 200
QY 57 PMGRIVEIYGPSSGKTTTLTQVIAAAQREGKTCAPDAEHALDPI----YARKLVGDID 112
Db 201 PDGNLIVIAARPSMGKTVLAENIARFALKQKQKAVHFQSYEMSAVELARRGMAECNIPMQ 260
QY 113 NL-----LCSPQDTGEQALE-----ICDAL-----ARSGAVDVIVVDVSV 146
Db 261 NLKTNLTQSDYANNPIYVSQAKWKFDVNCDDLNVDELCLFLAKEKKLTTGDLVLVDHL 320
QY 147 AALTPKA--EIEGEIGDSHMGLAARMQSMQMRKLAGNLKQSNLTLLIFINQIRMKIGVMFG 204
Db 321 -HIMPRAGRDEVAELGN-----ISRRLKNLAEL---NTPVVLVAQLNR-----G 361
QY 205 NPETTTGNALKFYASVR---LDIRRIGAVKEGENVV 238
Db 362 NTK-----QADKEPNMADIRGSAIEQDANII 388

RESULT 7

US-10-467-657-6400
; Sequence 6400, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON Spa
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 6400
; LENGTH: 453
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-6400

Query Match 6.5%; Score 109; DB 6; Length 453;

Best Local Similarity 23.5%; Pred. No. 0.36;
Matches 65; Conservative 43; Mismatches 91; Indels 78; Gaps 16;
QY 1 AIDENKOKALAAALGQ-IEKQFGKGSIMRLGEDRSMVDVET--ISTGSLSLDIALG-AGGL 56
Db 151 AVGDVAVKRETKTQGTVEDLIG-----GLDKRLDGVRFGLPTGLMKLD---GWTGGL 200
QY 57 PMGRIVEIYGPSSGKTTTLTQVIAAAQREGKTCAPDAEHALDPI----YARKLVGDID 112
Db 201 PDGNLIVIAARPSMGKTVLAENIARFALKQKQKAVHFQSYEMSAVELARRGMAECNIPMQ 260
QY 113 NL-----LCSPQDTGEQALE-----ICDAL-----ARSGAVDVIVVDVSV 146
Db 261 NLKTNLTQSDYANNPIYVSQAKWKFDVNCDDLNVDELCLFLAKEKKLTTGDLVLVDHL 320

```

; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqMan99, version 1.04
; SEQ ID NO 3626
; LENGTH: 453
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-3626

Query Match
Best Local Similarity 23.4%; Score 102; DB 6; Length 453;
Matches 68; Conservative 42; Mismatches 87; Indels 94; Gaps 17;

QY 2 IDENKQKA---LAAALGQIEKFGKSGIMRL-----GKDAVKRETFTGQTVVEDLIGGLDKRLDGVRGELPTG 189
Db 138 VAEKLSKADELAAA-----GKDAVKRETFTGQTVVEDLIGGLDKRLDGVRGELPTG 189

QY 44 SLSDIALG-AGGLPMGRIVEIYGPSSGKTTLTQLVIAAAQREGKTCAFIDAEHALDPI 102
Db 190 LMKLD---GMTGGLPDGNLIVTAARPSMGKTVLAENIARFALKQKGVHFSQVENSASVEL 246

QY 103 ----YARKLGVDIDNL-----LCQOPDTGEOALE-----ICDAL----- 132
Db 247 ARRGMAAECNIPMQLNKTGNLTQSDYANNPIYVSQAKWKPFVNCDLLNVDELCLFLAKEK 306

QY 133 ARSGAVDVIVDSVAALTPKA--EIEGEIGDSHMGLAARMQSMRKLKAGNLKQSNLTLLI 190
Db 307 KLTGTGLDLVVDHL-HIMPRAGRDEVAELGN-----ISRRLKNLAEL---NIPVV 353

QY 191 FINQIRMKIGVMFGNPETTTGNALKFYASVR---LDIRRIQAVKEGENVV 238
Db 354 LVAQLNR-----GNTK-----QADKRPNMADIRGSGAIEQADANII 388

RESULT 10
US-10-454-437-50
; Sequence 50, Application US/10454437
; Publication No. US2005027115A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Habernauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; TITLE OF INVENTION: INVOLVED IN HOMEOSTASIS AND ADAPTATION
; FILE REFERENCE: BGI-128CPCN
; CURRENT APPLICATION NUMBER: US/10/454,437
; CURRENT FILING DATE: 2003-06-13
; PRIOR APPLICATION NUMBER: US 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: DE 19931636.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932125.6
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932126.4
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932127.2
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932128.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932129.9
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: DE 19932226.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932920.6
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932922.2
; PRIOR FILING DATE: 1999-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 442
; SEQ ID NO 50
; LENGTH: 925
```

```

; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqMan99, version 1.04
; SEQ ID NO 3626
; LENGTH: 453
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-3626

Query Match
Best Local Similarity 23.4%; Score 102; DB 6; Length 453;
Matches 68; Conservative 42; Mismatches 87; Indels 94; Gaps 17;

QY 2 IDENKQKA---LAAALGQIEKFGKSGIMRL-----GKDAVKRETFTGQTVVEDLIGGLDKRLDGVRGELPTG 189
Db 138 VAEKLSKADELAAA-----GKDAVKRETFTGQTVVEDLIGGLDKRLDGVRGELPTG 189

QY 44 SLSDIALG-AGGLPMGRIVEIYGPSSGKTTLTQLVIAAAQREGKTCAFIDAEHALDPI 102
Db 190 LMKLD---GMTGGLPDGNLIVTAARPSMGKTVLAENIARFALKQKGVHFSQVENSASVEL 246

QY 103 ----YARKLGVDIDNL-----LCQOPDTGEOALE-----ICDAL----- 132
Db 247 ARRGMAAECNIPMQLNKTGNLTQSDYANNPIYVSQAKWKPFVNCDLLNVDELCLFLAKEK 306

QY 133 ARSGAVDVIVDSVAALTPKA--EIEGEIGDSHMGLAARMQSMRKLKAGNLKQSNLTLLI 190
Db 307 KLTGTGLDLVVDHL-HIMPRAGRDEVAELGN-----ISRRLKNLAEL---NIPVV 353

QY 191 FINQIRMKIGVMFGNPETTTGNALKFYASVR---LDIRRIQAVKEGENVV 238
Db 354 LVAQLNR-----GNTK-----QADKRPNMADIRGSGAIEQADANII 388

RESULT 10
US-10-454-437-50
; Sequence 50, Application US/10454437
; Publication No. US2005027115A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Habernauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; TITLE OF INVENTION: INVOLVED IN HOMEOSTASIS AND ADAPTATION
; FILE REFERENCE: BGI-128CPCN
; CURRENT APPLICATION NUMBER: US/10/454,437
; CURRENT FILING DATE: 2003-06-13
; PRIOR APPLICATION NUMBER: US 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: DE 19931636.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932125.6
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932126.4
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932127.2
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932128.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932129.9
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: DE 19932226.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932920.6
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932922.2
; PRIOR FILING DATE: 1999-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 442
; SEQ ID NO 50
; LENGTH: 925
```

```

; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqMan99, version 1.04
; SEQ ID NO 3626
; LENGTH: 453
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-3626

Query Match
Best Local Similarity 23.4%; Score 102; DB 6; Length 453;
Matches 68; Conservative 42; Mismatches 87; Indels 94; Gaps 17;

QY 2 IDENKQKA---LAAALGQIEKFGKSGIMRL-----GKDAVKRETFTGQTVVEDLIGGLDKRLDGVRGELPTG 189
Db 138 VAEKLSKADELAAA-----GKDAVKRETFTGQTVVEDLIGGLDKRLDGVRGELPTG 189

QY 44 SLSDIALG-AGGLPMGRIVEIYGPSSGKTTLTQLVIAAAQREGKTCAFIDAEHALDPI 102
Db 190 LMKLD---GMTGGLPDGNLIVTAARPSMGKTVLAENIARFALKQKGVHFSQVENSASVEL 246

QY 103 ----YARKLGVDIDNL-----LCQOPDTGEOALE-----ICDAL----- 132
Db 247 ARRGMAAECNIPMQLNKTGNLTQSDYANNPIYVSQAKWKPFVNCDLLNVDELCLFLAKEK 306

QY 133 ARSGAVDVIVDSVAALTPKA--EIEGEIGDSHMGLAARMQSMRKLKAGNLKQSNLTLLI 190
Db 307 KLTGTGLDLVVDHL-HIMPRAGRDEVAELGN-----ISRRLKNLAEL---NIPVV 353

QY 191 FINQIRMKIGVMFGNPETTTGNALKFYASVR---LDIRRIQAVKEGENVV 238
Db 354 LVAQLNR-----GNTK-----QADKRPNMADIRGSGAIEQADANII 388

RESULT 10
US-10-454-437-50
; Sequence 50, Application US/10454437
; Publication No. US2005027115A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Habernauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; TITLE OF INVENTION: INVOLVED IN HOMEOSTASIS AND ADAPTATION
; FILE REFERENCE: BGI-128CPCN
; CURRENT APPLICATION NUMBER: US/10/454,437
; CURRENT FILING DATE: 2003-06-13
; PRIOR APPLICATION NUMBER: US 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: DE 19931636.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932125.6
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932126.4
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932127.2
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932128.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932129.9
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: DE 19932226.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932920.6
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932922.2
; PRIOR FILING DATE: 1999-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 442
; SEQ ID NO 50
; LENGTH: 925
```

RESULT 13
US-11-082-389-132
; Sequence 132, Application US/11082389
; Publication No. US20050244935A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard


```

; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberkauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; TITLE OF INVENTION: INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE
; TITLE OF INVENTION: TRANSPORT
; FILE REFERENCE: BGI-131PCPCN
; CURRENT APPLICATION NUMBER: US/11/082,389
; PRIOR FILING DATE: 2005-03-16
; PRIOR APPLICATION NUMBER: US 09/603024
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 60/143262
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: US 60/151281
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19930487.4
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19930489.0
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19931549.3
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931550.7
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932134.5
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19941379.7
; PRIOR FILING DATE: 1999-08-31
; Remaining prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 446
; SEQ ID NO 132
; LENGTH: 365
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-11-082-389-132

Query Match 5.8%; Score 97; DB 7; Length 365;
Best Local Similarity 20.5%; Pred. No. 2.2;
Matches 67; Conservative 54; Mismatches 126; Indels 80; Gaps 12;

QY 7 QKALAAAL---GQIEKQFGKSGIMRLGEDRSMQVETISGSLSDLTALGAGLPMGRIVE 63
Db 114 QQLDASILDQPEIASDGSAPVIDVEASKSFETALHKVSLAVR-----KGDLLG 166

QY 64 IYGPSSGKTTLTQVIAAAQREGKTCFAIDAEHALDPIYARKLGVDIDNLLCSQPDGTGE 123
Db 167 IVGGSGSGKTTL-LKLLA-----GLD-----KPTTG- 191

QY 124 QALEICDALARSGAVDVIVVDSVAALTPKAEIEGIDSHVGLAARMMSQMRKLAGNLK 183
Db 192 -----TVAVTGVQVHFQDPOSSLNPRMKIDIVAEPLLGWNA-----AE 231

QY 184 QSNLTLLIFINQIRKMGKMFNPGNETTTGGNALKFYASVRDIDIRIGAVKGENVVGSETR 243
Db 232 KTRFRAEVITQVGLSPDLDYRPHFSGGQ-----QRISARALAIPKA--ILLADEP 283

QY 244 KVVKVKIAFFQKQAEFQILYGEINIFYGELVDLVKVEKLEKAGAWSYKGEKIGOGKA 303
Db 284 VSALDVSVRKQVLDLQQLVEYGYTLVFSVSHDLAVVRHLC--TTVWVMEQGRVLEGGPI 341

QY 304 NATAWLKDNPET--AKIEKKVRELL 328
Db 342 DS-----VYDHPQETTKELLDVAPRLSL 365

```

```

RESULT 14
US-11-213-368-14
; Sequence 14, Application US/11213368
; Publication No. US20060003936A1
; GENERAL INFORMATION:
; APPLICANT: MayerKuckuk, Phillip
; APPLICANT: Banerjee, Debabrata

```

```

; APPLICANT: Bertino, Joseph R.
; TITLE OF INVENTION: Method for Modulating the Production of a Selected
; TITLE OF INVENTION: Protein In
; TITLE OF INVENTION: Vivo
; FILE REFERENCE: MSK.P-053
; CURRENT APPLICATION NUMBER: US/11/213,368
; CURRENT FILING DATE: 2005-08-26
; PRIOR APPLICATION NUMBER: US/10/421,285
; PRIOR FILING DATE: 2003-04-22
; PRIOR APPLICATION NUMBER: US 60/375,250
; PRIOR FILING DATE: 2002-04-22
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 14
; LENGTH: 376
; TYPE: PRT
; ORGANISM: herpes virus
US-11-213-368-14

Query Match 5.8%; Score 97; DB 7; Length 376;
Best Local Similarity 23.6%; Pred. No. 2.3;
Matches 53; Conservative 30; Mismatches 80; Indels 62; Gaps 11;

QY 58 MGRIVEIY--GPSSGKTTLTQVIAAAQREGKTCFAIDAEHALDPI-YARKLGVD--ID 112
Db 46 MPTLLRVYIDGPHGMGKTTTTQLLVALGSRD-----DIVVDPDMTYQVVLGASSETIA 98

QY 113 NLLCSQP-----DTGQALEICDALARSGAVDVIVVDSVAALTPKAEIEGIDSHM- 164
Db 99 NIYTTQHRLDQGEISAGDAAVMTSAQITMG-MPYAVTDAVLA-----PHIGGEAGSSHAP 153

QY 165 -----GLAARMMSQMRKLAGNLKQSNLTLLIFINQIRKMGKMFNPGNETTTGGN 213
Db 154 PPALTLLIFDRHPFAALLCYPAARYLMGSM-TPQAVLAFVALLI-----PPTLPUGTN 202

QY 214 -----ALKFYASVRDIDIRIGAVKGENVVGSETR 243
Db 203 IVLGALPEDRHDRLAKRQPCGERLDLMLAAIRRVYGLLANTVR 247

```

```

RESULT 15
US-11-082-389-130
; Sequence 130, Application US/11082389
; Publication No. US20050244935A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberkauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; TITLE OF INVENTION: INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE
; TITLE OF INVENTION: TRANSPORT
; FILE REFERENCE: BGI-131PCPCN
; CURRENT APPLICATION NUMBER: US/11/082,389
; CURRENT FILING DATE: 2005-03-16
; PRIOR APPLICATION NUMBER: US 09/603024
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 60/143262
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: US 60/151281
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19930487.4
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19930489.0
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19931549.3
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931550.7
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932134.5

```

```
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19941379.7
; PRIOR FILING DATE: 1999-08-31
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 446
; SEQ ID NO 130
; LENGTH: 479
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
; US-11-082-389-130

Query Match      5.8%; Score 97; DB 7; Length 479;
Best Local Similarity 20.5%; Pred. No. 3.2;
Matches 67; Conservative 54; Mismatches 126; Indels 80; Gaps 12;

QY      7 QKALAAAL---GQIEKQFGKSGIMRLGEDRSMDVETISTGSLSLDIALGAGGLPMGRIVE 63
Db      228 QQLLDASILDQPEIASDSGAPVVIDVEEASKSFKETTALHKVSLAVR-----KGDLLG 280

QY      64 IYGPESGKTTLTQVIAAAQREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPDGTGE 123
Db      281 IVGGSGGKTTL-LKLIA-----GLD-----KPTTG- 305

QY      124 QALEICDALARSGAVDIVVDSVAALTPKAEIEGEIGDSHMGLAARMMSQAMRKLAGNLK 183
Db      306 -----TVAVTGGVQMVFDQPOSSLNPRMKIKDIVAEPLLGWNA-----AE 345

QY      184 QSNTHLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGSETR 243
Db      346 KTRVAEVITQVGLSPDLDRYPHEFSGQ-----RQISIAALAIKPA--ILLADEP 397

QY      244 VKVVKNKIAAPFKQAEFQILYGEGINFYGELVDLGKKEKLIBKAGAWSYKGEKIQGQKA 303
Db      398 VSALDVSVRKQVLDLLQQLVEEYGITLVFVSHDLAVVRHLC--TTVWVMEQGRVLEQGP 455

QY      304 NATAWLKONPET--AKEIEKKVRELLJ 328
Db      456 DS---VYDHPQTEYTKELLDVAPRLSL 479
```

Search completed: February 16, 2006, 01:45:01
Job time : 11.5 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: February 16, 2006, 01:14:10 ; Search time 132 Seconds
(without alignments)
1115.090 Million cell updates/sec

Title: US-10-733-782-3
Perfect score: 1678
Sequence: 1 AIDENKQKALAAALGQIEKQ.....AKBIEKKVRELLSNPNSTP 335

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21:.*
1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*
9: Geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1678	100.0	335	8	ADR16430 E. coli R
2	1674	99.8	335	8	ADR16428 E. coli R
3	1674	99.8	352	8	ADR16432 E. coli w
4	1674	99.8	353	6	ABU15395 Protein e
5	1674	99.8	358	3	AAV68827 Amino aci
6	1674	99.8	358	3	AAV68829 Amino aci
7	1674	99.8	358	8	ADQ87824 E. coli r
8	1674	99.8	358	8	ADQ87822 E. coli r
9	1674	99.8	358	8	ADQ87821 E. coli r
10	1673	99.7	358	3	AAV68831 Amino aci
11	1670	99.5	358	2	AAW64213 New minsh
12	1669	99.5	358	3	AAV68828 Amino aci
13	1669	99.5	358	8	ADQ87823 E. coli r
14	1665	99.2	358	8	ADQ87826 E. coli r
15	1664	99.2	358	2	AAW64215 Hyperreco
16	1661	99.0	358	2	AAW64214 Hyperreco
17	1658	98.8	358	2	AAW64217 Hyperreco
18	1658	98.8	358	2	AAW64218 Hyperreco
19	1656	98.7	358	3	AAV68832 Amino aci
20	1656	98.7	358	8	ADQ87827 E. coli r
21	1655	98.6	358	3	AAV68830 Amino aci
22	1655	98.6	358	8	ADQ87825 E. coli r
23	1639	97.7	358	2	AAW64216 Hyperreco
24	1637	97.6	353	6	ABU48024 Protein e

25	1609	95.9	352	6	ABU31601 Protein e
26	1609	95.9	356	7	ABO63433 Klebsiell
27	1589	94.7	352	6	ABU27708 Protein e
28	1543	92.0	356	6	ABM67793 Phototrab
29	1530	91.2	356	6	ABU50600 Protein e
30	1500	89.4	355	6	ABU40669 Protein e
31	1500	89.4	367	7	ADF05150 Bacterial
32	1463	87.2	351	6	ABU44971 Protein e
33	1432	85.3	412	6	ABU49151 Protein e
34	1349	80.4	354	6	ABU39491 Protein e
35	1328	79.1	354	6	ABU30311 Protein e
36	1269	75.6	398	7	ABO72381 Pseudomon
37	1268	75.6	346	6	ABU38581 Protein e
38	1257	74.9	355	6	ABU40018 Protein e
39	1255	74.8	349	6	ABU16844 Protein e
40	1255	74.8	355	6	ADA36276 Acinetoba
41	1255	74.8	369	6	ABU41878 Protein e
42	1247.5	74.3	366	8	ADL05025 M. catarr
43	1242.5	74.0	349	6	ABU35468 Protein e
44	1238.5	73.8	342	8	ADR21929 Methylph
45	1237.5	73.7	348	6	ABU33125 Protein e

ALIGNMENTS

RESULT 1
ADR16430
ID ADR16430 standard; protein; 335 AA.
XX
AC ADR16430;
XX
DT 04-NOV-2004 (first entry)
XX
DE E. coli RecA mutant protein #2.
XX
KW RecA; DNA binding protein; ssDNA binding protein; SSB;
KW DNA strand exchange reaction; site-specific cleavage; mutant; muten.
XX
OS Escherichia coli.
XX
FH Key Location/Qualifiers
FT Misc-difference 38
FT /note= "Wild-type Glu substituted by Lys"
XX
XX US2004157248-A1.
XX
PD 12-AUG-2004.
XX
XX 11-DEC-2003; 2003US-00733782.
XX
PR 12-DEC-2002; 2002US-0432758P.
XX
XX (COXM/) COX M M.
XX (LUSE/) LUSSETTI S L.
XX (EGGL/) EGGLER A L.
XX (HARU/) HARUTA N.
XX
XX Cox MM, Lusettti SL, Egglar AL, Haruta N;
XX
XX WPI; 2004-580265/56.
XX N-PSDB; ADR16431.
XX
XX New RecA mutant proteins comprising a single mutation or a double
XX mutation, useful for catalyzing homologous DNA pairing and DNA strand
XX exchange reactions in an in vitro or in vivo environment.
XX
XX Claim 13; SEQ ID NO 3; 34pp; English.
XX
XX The invention relates to an isolated RecA mutant protein, where the
XX protein is either a single mutant RecA protein comprising a deletion of
XX amino acid residues from the carboxyl terminus, or a double mutant RecA
XX protein comprising a deletion of amino acid residues from the carboxyl

terminus and an amino acid change from a glutamate to a basic amino acid.
The invention also relates to a polynucleotide sequence that encodes the
Reca mutant protein, a method of catalysing an in vitro homologous DNA
pairing and DNA strand exchange reactions comprising providing an amount
of the RecA mutant protein and a method of increasing recombination
efficiency of homologous DNA pairing and DNA strand exchange reactions in
a cell comprising supplying to the cell an amount of the RecA mutant
protein. The RecA mutant protein comprises an enhanced capacity to
displace a DNA binding protein as compared to wild-type RecA. The DNA
binding protein is the ssDNA binding protein from *Escherichia coli* (SSB).
It also comprises enhanced binding to DNA during a DNA strand exchange
reaction as compared to wild-type RecA. The composition and methods are
useful for catalysing homologous DNA pairing and DNA strand exchange
reactions in an in vitro or in vivo environment. These may be used in
promoting in vitro alterations of genes to permit the rapid construction
of desired gene mutants for industrial and pharmaceutical purposes. The
mutant proteins may also be used in targeting the site-specific cleavage
of small and large DNAs, or as a basis for the design and construction of
tiny electronic circuits based on DNA. This sequence represents a RecA
mutant protein of the invention.

XX SQ Sequence 335 AA;
Query Match 100.0%; Score 1678; DB 8; Length 335;
Best Local Similarity 100.0%; Pred. No. 2.8e-155;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMDVKTISTGSLSLDIALGAGGLPMGR 60
DB 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMDVKTISTGSLSLDIALGAGGLPMGR 60
QY 61 IVEIYGPESSGKTTTLQVIAAAREGKTCAPIDAEHALDPIYARKLGVDDINLLCSQPD 120
DB 61 IVEIYGPESSGKTTTLQVIAAAREGKTCAPIDAEHALDPIYARKLGVDDINLLCSQPD 120
QY 121 TGEQALEICDALARSGAVDVIIVDSVAALTTPKAEIEGEGISHMGLAARMMSQAMRKLAG 180
DB 121 TGEQALEICDALARSGAVDVIIVDSVAALTTPKAEIEGEGISHMGLAARMMSQAMRKLAG 180
QY 181 NLKQNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVG 240
DB 181 NLKQNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVG 240
QY 241 ETRVVKVKNKIAAPFKQAEFOILYEGEGINFYGELVDLGVKELIEKAGAWSYKGEKIQ 300
DB 241 ETRVVKVKNKIAAPFKQAEFOILYEGEGINFYGELVDLGVKELIEKAGAWSYKGEKIQ 300
QY 301 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 335
DB 301 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 335

RESULT 2

ADRI16428
ID ADRI16428 standard; protein; 335 AA.
AC ADRI16428;
XX
XX 04-NOV-2004 (first entry)
XX
XX E. coli RecA mutant protein #1.
XX
XX RecA; DNA binding protein; ssDNA binding protein; SSB;
KW DNA strand exchange reaction; site-specific cleavage; mutant; muten.
XX
XX *Escherichia coli*.
XX
XX US2004157248-A1.
XX
XX 12-AUG-2004.
XX
XX 11-DEC-2003; 2003US-00733782.
XX

PR 12-DEC-2002; 2002US-0432758P.

XX (COXM/) COX M M.
PA (LUSE/) LUSSETTI S L.
PA (EGL/) EGLER A L.
XX (HARU/) HARUTA N.
PI Cox MM, Lusetti SL, Eggler AL, Haruta N;
XX WPI; 2004-580265/56.
DR N-PSDB; ADR16429.

XX New RecA mutant proteins comprising a single mutation or a double
mutation, useful for catalysing homologous DNA pairing and DNA strand
exchange reactions in an in vitro or in vivo environment.

Claim 2; SEQ ID NO 1; 34pp; English.

The invention relates to an isolated RecA mutant protein, where the
protein is either a single mutant RecA protein comprising a deletion of
amino acid residues from the carboxyl terminus, or a double mutant RecA
protein comprising a deletion of amino acid residues from the carboxyl
terminus and an amino acid change from a glutamate to a basic amino acid.
The invention also relates to a polynucleotide sequence that encodes the
Reca mutant protein, a method of catalysing an in vitro homologous DNA
pairing and DNA strand exchange reactions comprising providing an amount
of the RecA mutant protein and a method of increasing recombination
efficiency of homologous DNA pairing and DNA strand exchange reactions in
a cell comprising supplying to the cell an amount of the RecA mutant
protein. The RecA mutant protein comprises an enhanced capacity to
displace a DNA binding protein as compared to wild-type RecA. The DNA
binding protein is the ssDNA binding protein from *Escherichia coli* (SSB).
It also comprises enhanced binding to DNA during a DNA strand exchange
reaction as compared to wild-type RecA. The composition and methods are
useful for catalysing homologous DNA pairing and DNA strand exchange
reactions in an in vitro or in vivo environment. These may be used in
promoting in vitro alterations of genes to permit the rapid construction
of desired gene mutants for industrial and pharmaceutical purposes. The
mutant proteins may also be used in targeting the site-specific cleavage
of small and large DNAs, or as a basis for the design and construction of
tiny electronic circuits based on DNA. This sequence represents a RecA
mutant protein of the invention.

XX SQ Sequence 335 AA;

Query Match 99.8%; Score 1674; DB 8; Length 335;
Best Local Similarity 99.7%; Pred. No. 7e-155;
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMDVKTISTGSLSLDIALGAGGLPMGR 60
DB 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMDVKTISTGSLSLDIALGAGGLPMGR 60
QY 61 IVEIYGPESSGKTTTLQVIAAAREGKTCAPIDAEHALDPIYARKLGVDDINLLCSQPD 120
DB 61 IVEIYGPESSGKTTTLQVIAAAREGKTCAPIDAEHALDPIYARKLGVDDINLLCSQPD 120
QY 121 TGEQALEICDALARSGAVDVIIVDSVAALTTPKAEIEGEGISHMGLAARMMSQAMRKLAG 180
DB 121 TGEQALEICDALARSGAVDVIIVDSVAALTTPKAEIEGEGISHMGLAARMMSQAMRKLAG 180
QY 181 NLKQNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVG 240
DB 181 NLKQNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVG 240
QY 241 ETRVVKVKNKIAAPFKQAEFOILYEGEGINFYGELVDLGVKELIEKAGAWSYKGEKIQ 300
DB 241 ETRVVKVKNKIAAPFKQAEFOILYEGEGINFYGELVDLGVKELIEKAGAWSYKGEKIQ 300
QY 301 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 335
DB 301 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 335

RESULT 3	
ID	ADRI16432
XX	ADRI16432 standard; protein; 352 AA.
AC	ADRI16432;
XX	
DT	04-NOV-2004 (first entry)
XX	
DE	E. coli wild-type RecA protein.
XX	
KW	RecA; DNA binding protein; ssDNA binding protein; SSB;
KW	DNA strand exchange reaction; site-specific cleavage.
XX	
OS	Escherichia coli.
XX	
PN	US2004157248-A1.
XX	
PD	12-AUG-2004.
XX	
PF	11-DEC-2003; 2003US-00733782.
XX	
PR	12-DEC-2002; 2002US-0432758P.
XX	
PA	(COXM//) COX M M.
FA	(LUSE//) LUSSETTI S L.
FA	(EGGL//) EGGLE A L.
PA	(HARU//) HARUTA N.
XX	
PI	Cox MM, Lusetti SL, Eggler AL, Haruta N;
XX	
DR	WPI; 2004-580265/56.
XX	
PT	New RecA mutant proteins comprising a single mutation or a double
PT	mutation, useful for catalyzing homologous DNA pairing and DNA strand
PT	exchange reactions in an in vitro or in vivo environment.
XX	
FS	Disclosure; SEQ ID NO 5; 34pp; English.
XX	
CC	The invention relates to an isolated RecA mutant protein, where the
CC	protein is either a single mutant RecA protein comprising a deletion of
CC	amino acid residues from the carboxyl terminus, or a double mutant RecA
CC	protein comprising a deletion of amino acid residues from the carboxyl
CC	terminus and an amino acid change from a glutamate to a basic amino acid.
CC	The invention also relates to a polynucleotide sequence that encodes the
CC	RecA mutant protein, a method of catalyzing an in vitro homologous DNA
CC	pairing and DNA strand exchange reactions comprising providing an amount
CC	of the RecA mutant protein and a method of increasing recombination
CC	efficiency of homologous DNA pairing and DNA strand exchange reactions in
CC	a cell comprising supplying to the cell an amount of the RecA mutant
CC	protein. The RecA mutant protein comprises an enhanced capacity to
CC	displace a DNA binding protein as compared to wild-type RecA. The DNA
CC	binding protein is the ssDNA binding protein from Escherichia coli (SSB).
CC	It also comprises enhanced binding to DNA during a DNA strand exchange
CC	reaction as compared to wild-type RecA. The composition and methods are
CC	useful for catalyzing homologous DNA pairing and DNA strand exchange
CC	reactions in an in vitro or in vivo environment. These may be used in
CC	promoting in vitro alterations of genes to permit the rapid construction
CC	of desired gene mutants for industrial and pharmaceutical purposes. The
CC	mutant proteins may also be used in targeting the site-specific cleavage
CC	of small and large DNAs, or as a basis for the design and construction of
CC	tiny electronic circuits based on DNA. This sequence represents the wild-
CC	type RecA protein of the invention.
XX	
SQ	Sequence 352 AA;
XX	
Query Match	99.8%; Score 1674; DB 8; Length 352;
Best Local Similarity	99.7%; Pred. NO. 7.5e-155;
Matches 334; Conservative	1; Mismatches 0; Indels 0; Gaps 0;
Qy	1 AIDENKQKALAAALGQIEKQFGKSIIMRLGDRSDMDVKTISTGSLDIALGAGGLPMGR 60
Db	1 AIDENKQKALAAALGQIEKQFGKSIIMRLGDRSDMDVKTISTGSLDIALGAGGLPMGR 60

QY	61	IVEIYGPESSGKTTLTLLQVIAAAQREGKTCAFIDAEHALDPIYARKLGVYDIDNLLCSQPD	120
Db	61	IVEIYGPESSGKTTLTLLQVIAAAQREGKTCAFIDAEHALDPIYARKLGVYDIDNLLCSQPD	120
QY	121	TGEQALEICDALARSGAVDVIVDSVAALTPKAEIEGEIGDSHMGGLAARMMSQAMRKLAG	180
Db	121	TGEQALEICDALARSGAVDVIVDSVAALTPKAEIEGEIGDSHMGGLAARMMSQAMRKLAG	180
QY	181	NLKOSNTLLIFINQIRMKIGVMFGNPETTTGNGALKFYASVRLDIRRGAVKEGENVVGS	240
Db	181	NLKOSNTLLIFINQIRMKIGVMFGNPETTTGNGALKFYASVRLDIRRGAVKEGENVVGS	240
QY	241	ETRVKVVKNKIAAPFKQAEFQILYGEINFGYELVDLGVKBLIEKAGAWSYKGEKIQG	300
Db	241	ETRVKVVKNKIAAPFKQAEFQILYGEINFGYELVDLGVKBLIEKAGAWSYKGEKIQG	300
QY	301	GKANATAWLKONPETAKEIEKKVRELLLSNPNSTP	335
Db	301	GKANATAWLKONPETAKEIEKKVRELLLSNPNSTP	335

RESULT 4

ABUL5395

ID ABUL5395 standard; protein; 353 AA.

XX

AC ABUL5395;

XX

DT 19-JUN-2003 (first entry)

XX

DE Protein encoded by Prokaryotic essential gene #922.

XX

KW Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX

OS Escherichia coli.

XX

PN WO200277183-A2.

XX

PD 03-OCT-2002.

XX

PF 21-MAR-2002; 2002WO-US009107.

XX

PR 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

XX

(ELIT-) ELITRA PHARM INC.

PA Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX

WPI; 2003-029926/02.

DR N-ESDB; ACA19265.

XX

New antisense nucleic acids, useful for identifying proteins or screening

PT for homologous nucleic acids required for cellular proliferation to

PT isolate candidate molecules for rational drug discovery programs.

XX

Claim 25; SEQ ID NO 43319; 1766pp; English.

XX

The invention relates to an isolated nucleic acid comprising any one of

CC the 6213 antisense sequences given in the specification where expression

CC of the nucleic acid inhibits proliferation of a cell. Also included are:

CC (1) a vector comprising a promoter operably linked to the nucleic acid

CC encoding a polypeptide whose expression is inhibited by the antisense

CC nucleic acid; (2) a host cell containing the vector; (3) an isolated

CC polypeptide or its fragment whose expression is inhibited by the

CC antisense nucleic acid; (4) an antibody capable of specifically binding

CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular

CC proliferation or the activity of a gene in an operon required for

CC proliferation; (7) identifying a compound that influences the activity of

CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 353 AA;

Query Match 99.8%; Score 1674; DB 6; Length 353;
 Best Local Similarity 99.7%; Pred. No. 7.6e-155;
 Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AIDENKQKALAAALGOIEKQFGKSGIMRLGEDRSMVDVTISTGSLDLALGAGGLPMGR 60
 DB 2 AIDENKQKALAAALGOIEKQFGKSGIMRLGEDRSMVDVTISTGSLDLALGAGGLPMGR 61

QY 61 IVEIYGPESSGKTTTLQVIAAAOQREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
 DB 62 IVEIYGPESSGKTTTLQVIAAAOQREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 121

QY 121 TGEQALEICDALARSGAVDVIWVDSVAALTPKAEIEGEGIGDSHMGLAARMMSQAMRKLKAG 180
 DB 122 TGEQALEICDALARSGAVDVIWVDSVAALTPKAEIEGEGIGDSHMGLAARMMSQAMRKLKAG 181

QY 181 NLKQSNLTLLIFINQIRMKIGVFNFGNPTTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
 DB 182 NLKQSNLTLLIFINQIRMKIGVFNFGNPTTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 241

QY 241 ETRVKKVKNKIAAPFKQAEFQILYGEINFGYELVDLGVEKLEIKAGAWSYKGEKIGQ 300
 DB 242 ETRVKKVKNKIAAPFKQAEFQILYGEINFGYELVDLGVEKLEIKAGAWSYKGEKIGQ 301

QY 301 GKANATAWLKONPETAKEIEKKVRELLLSNPNSTP 335
 DB 302 GKANATAWLKONPETAKEIEKKVRELLLSNPNSTP 336

RESULT 5

AAV68827

ID AAY68827 standard; protein; 358 AA.

XX AC AAY68827;

XX DT 16-MAY-2000 (first entry)

XX DE Amino acid sequence of a wildtype recA protein.

XX KW recA; hyperrecombinogenic variant; male gamete; female gamete;
 XX KW heat tolerance; ethanol production; ethanol tolerance; metabolite; taxol;
 XX KW cyclosporin A; erythromycin; meiosis.

XX OS Escherichia coli.

XX FN WO200004190-A1.

XX PD 27-JAN-2000.

XX PF 15-JUL-1999; 99WO-US015972.

XX 15-JUL-1998; 98US-00116188.

XX PA (MAXY-) MAXYGEN INC.

XX PI Del Cardayre S, Tobin M, Stemmer WPC, Ness JE, Minshull J;

XX PI Patten PA, Subramanian V, Castle LA, Krebber CM, Bass S, Zhang Y;

XX PI Cox T, Huisman G, Yuan L, Affholter JA;

XX PI WPI; 2000-182446/16.

XX DR N-PSDB; AAZ60609.

XX PT Evolution of whole cells and organism by iterative cycles of

XX PT recombination and selection and screening for acquisition of desired

XX PT properties.

XX PS Example 1; Fig 13; 197pp; English.

CC The present sequence represents the recA protein. The wild type recA
 CC protein was used to produce hyperrecombinogenic variants (see AAY68828-
 CC 32), using the method of the invention. The specification describes a
 CC method for producing a library of diverse multicellular organisms using
 CC pools of male and female gametes. At least one of the male pool or female
 CC strains comprises a number of different gametes derived from different
 CC pools of a species or of a different species. The viable organisms
 CC produced from the fertilized gametes are repeatedly crossed to produce a
 CC library of diverse organisms, which are selected for a desired trait or
 CC property. The methods can be used to evolve cells to acquire a desired
 CC property such as heat tolerance, ethanol production or tolerance, acid,
 CC improved production and maintenance of enzyme cofactors or NAD(P)H and
 CC improved glucose transport. The desired property may be expression of a
 CC protein or primary or secondary metabolite. Alternatively the desired
 CC property is secretion of a protein or secondary metabolite, chosen from
 CC taxol, cyclosporin A and erythromycin. The desired property may be a
 CC capacity for meiosis or compatibility to form a heterokaryon with another
 CC strain

XX SQ Sequence 358 AA;

Query Match 99.8%; Score 1674; DB 3; Length 358;
 Best Local Similarity 99.7%; Pred. No. 7.7e-155;
 Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AIDENKQKALAAALGOIEKQFGKSGIMRLGEDRSMVDVTISTGSLDLALGAGGLPMGR 60
 DB 7 AIDENKQKALAAALGOIEKQFGKSGIMRLGEDRSMVDVTISTGSLDLALGAGGLPMGR 66

QY 61 IVEIYGPESSGKTTTLQVIAAAOQREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
 DB 67 IVEIYGPESSGKTTTLQVIAAAOQREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 126

QY 121 TGEQALEICDALARSGAVDVIWVDSVAALTPKAEIEGEGIGDSHMGLAARMMSQAMRKLKAG 180
 DB 127 TGEQALEICDALARSGAVDVIWVDSVAALTPKAEIEGEGIGDSHMGLAARMMSQAMRKLKAG 186

QY 181 NLKQSNLTLLIFINQIRMKIGVFNFGNPTTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
 DB 187 NLKQSNLTLLIFINQIRMKIGVFNFGNPTTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 246

QY 241 ETRVKKVKNKIAAPFKQAEFQILYGEINFGYELVDLGVEKLEIKAGAWSYKGEKIGQ 300
 DB 247 ETRVKKVKNKIAAPFKQAEFQILYGEINFGYELVDLGVEKLEIKAGAWSYKGEKIGQ 306

QY 301 GKANATAWLKONPETAKEIEKKVRELLLSNPNSTP 335
 DB 307 GKANATAWLKONPETAKEIEKKVRELLLSNPNSTP 341

Query Match

Best Local Similarity 99.7%; Pred. No. 7.7e-155;

Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AIDENKQKALAAALGOIEKQFGKSGIMRLGEDRSMVDVTISTGSLDLALGAGGLPMGR 60

DB 7 AIDENKQKALAAALGOIEKQFGKSGIMRLGEDRSMVDVTISTGSLDLALGAGGLPMGR 66

QY 61 IVEIYGPESSGKTTTLQVIAAAOQREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 120

DB 67 IVEIYGPESSGKTTTLQVIAAAOQREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 126

QY 121 TGEQALEICDALARSGAVDVIWVDSVAALTPKAEIEGEGIGDSHMGLAARMMSQAMRKLKAG 180

DB 127 TGEQALEICDALARSGAVDVIWVDSVAALTPKAEIEGEGIGDSHMGLAARMMSQAMRKLKAG 186

QY 181 NLKQSNLTLLIFINQIRMKIGVFNFGNPTTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240

DB 187 NLKQSNLTLLIFINQIRMKIGVFNFGNPTTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 246

QY 241 ETRVKKVKNKIAAPFKQAEFQILYGEINFGYELVDLGVEKLEIKAGAWSYKGEKIGQ 300

DB 247 ETRVKKVKNKIAAPFKQAEFQILYGEINFGYELVDLGVEKLEIKAGAWSYKGEKIGQ 306

QY 301 GKANATAWLKONPETAKEIEKKVRELLLSNPNSTP 335

DB 307 GKANATAWLKONPETAKEIEKKVRELLLSNPNSTP 341

RESULT 6

AAV68829

ID AAY68829 standard; protein; 358 AA.

XX AC AAY68829;


```

XX 16-MAY-2000 (first entry)
XX Amino acid sequence of a hyperrecombinogenic recA protein clone 4.
XX recA; hyperrecombinogenic variant; male gamete; female gamete;
XX taxol; ethanol production; ethanol tolerance; metabolite; taxol;
XX cyclosporin A; erythromycin; meiosis.
XX Synthetic.
XX Escherichia coli.
XX WO200004190-A1.
XX 27-JAN-2000.
XX 15-JUL-1999; 99WO-US015972.
XX 15-JUL-1998; 98US-00116188.
XX (MAXY-) MAXYGEN INC.
XX Del Cardayre S, Tobin M, Stemmer WPC, Ness JE, Minshull J;
XX Patten PA, Subramanian V, Castle LA, Krebber CM, Bass S, Zhang Y;
XX Cox T, Huisman G, Yuan L, Affholter JA;
XX WPI; 2000-182446/16.
XX N-PSDB; NAE60611.
XX Evolution of whole cells and organism by iterative cycles of
XX recombination and selection and screening for acquisition of desired
XX properties.
XX Example 1; Fig 13; 197pp; English.
XX The present sequence represents a hyperrecombinogenic recA protein. The
XX wild type recA protein (AAV68827) was used to produce hyperrecombinogenic
XX variants, using the method of the invention. The specification describes
XX a method for producing a library of diverse multicellular organisms using
XX pools of male and female gametes. At least one of the male pool or female
XX pools comprises a number of different gametes derived from different
XX strains of a species or of a different species. The viable organisms
XX produced from the fertilized gametes are repeatedly crossed to produce a
XX library of diverse organisms, which are selected for a desired trait or
XX property. The methods can be used to evolve cells to acquire a desired
XX property such as heat tolerance, ethanol production or tolerance, acid,
XX improved production and maintenance of enzyme cofactors or NAD(P)H and
XX improved glucose transport. The desired property may be expression of a
XX protein or primary or secondary metabolite. Alternatively the desired
XX property is secretion of a protein or secondary metabolite, chosen from
XX taxol, cyclosporin A and erythromycin. The desired property may be a
XX capacity for meiosis or compatibility to form a heterokaryon with another
XX strain
XX SQ Sequence 358 AA;
Query Match 99.8%; Score 1674; DB 3; Length 358;
Best Local Similarity 99.7%; Pred. No. 7.7e-155;
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGSDRSMVDVETISTGSLDLIALGAGGLPMGR 60
DB 7 AIDENKQKALAAALGQIEKQFGKGSIMRLGSDRSMVDVETISTGSLDLIALGAGGLPMGR 66
QY 61 IVEIYGESSGKTTLTQVIAAAREGKTCAFIDAEHALDPYARKLGVDIDNLCSQPD 120
DB 67 IVEIYGESSGKTTLTQVIAAAREGKTCAFIDAEHALDPYARKLGVDIDNLCSQPD 126
QY 121 TGEQALIEICDALARSGAVDVIVDVSVAALTPKABIEIGIGDSHMLAARWMSQAMRKLKAG 180
DB 127 TGEQALIEICDALARSGAVDVIVDVSVAALTPKABIEIGIGDSHMLAARWMSQAMRKLKAG 186
QY 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240

```

```

DB 187 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 246
QY 241 ETRVKVVKNTIAAPFKQAEFOILYGEINFGYELVDLGKVKLIEKAGAWSYKGEKIGQ 300
DB 247 ETRVKVVKNTIAAPFKQAEFOILYGEINFGYELVDLGKVKLIEKAGAWSYKGEKIGQ 306
QY 301 GKANATAWLKDNPETAKEIEKKVRELLSNPNSTP 335
DB 307 GKANATAWLKDNPETAKEIEKKVRELLSNPNSTP 341
RESULT 7
ADQ87824
ID ADQ87824 standard; protein; 358 AA.
XX AC ADQ87824;
XX DT 09-SEP-2004 (first entry)
XX DE E. coli recA protein clone #3.
XX KW RecA; protoplast formation; gamete; heat tolerance; ethanol production;
XX ethanol tolerance; enzyme cofactor; NAD(P)H; glucose transport; meiosis;
XX heterokaryon; taxol; cyclosporin A; erythromycin; viral infection;
XX plant genome shuffling; micropore manipulation; reiterative pooling.
XX OS Escherichia coli.
XX FN AU2004200501-A1.
XX PD 04-MAR-2004.
XX PF 09-FEB-2004; 2004AU-00200501.
XX PR 09-FEB-2004; 2004AU-00200501.
XX PA (MAXY-) MAXYGEN INC.
XX PI Yuan L, Cox T, Bass S, Castle LA, Patten PA, Ness JE, Tobin M;
XX Affholter JA, Huisman G, Zhang Y, Krebber CM, Subramanian V;
XX PI Minshull J, Stemmer WPC, Del Cardayre S;
XX DR WPI; 2004-507924/49.
XX DR N-PSDB; ADQ87817.
XX PT Evolving cells to acquire a desired property, by forming protoplasts of
XX different cells, fusing protoplasts to form hybrid protoplasts (HP),
XX producing regenerated cells, forming additional HP and producing
XX additional regenerated cells.
XX PS Disclosure; Fig 13; 196pp; English.
XX The invention relates to a method of evolving cells to acquire a desired
XX property by forming protoplasts of different cells, fusing the
XX protoplasts to form hybrid protoplasts, incubating the hybrid protoplasts
XX to produce regenerated cells, repeatedly forming protoplasts from
XX regenerated cells, fusing protoplasts to form hybrid protoplasts in which
XX genomes from protoplasts recombine to form additional hybrid genomes and
XX incubating additional hybrid protoplasts for producing additional
XX regenerated cells. The invention also relates to a method of producing a
XX library of diverse multicellular organisms involving providing a pool of
XX male gametes and a pool of female gametes, where one of the male pool or
XX the female pool comprises several different gametes derived from
XX different strains of a species or different species and the male gametes
XX fertilise the female gametes, permitting at least a portion of the
XX resulting fertilised gametes to grow into reproductively viable
XX organisms, repeatedly crossing the reproductively viable organisms to
XX produce a library of diverse organisms and selecting the library for a
XX desired trait or property. The methods are useful for evolving cells to
XX acquire desired properties such as heat tolerance, ethanol production,
XX ethanol tolerance, improved production and maintenance of enzyme
XX cofactors, improved production and maintenance of NAD(P)H and improved

```

CC glucose transport. The desired property is the expression of a protein,
CC primary metabolite or secondary metabolite, the secretion of a protein or
CC heterokaryon with another strain. The secondary metabolite is chosen from
CC taxol, cyclosporin A and erythromycin. The method is also useful for
CC acquisition of desired properties such as enhanced recombination, or
CC gene copy number or gene reductivity and capacity for expression and/or
CC secretion of proteins or secondary metabolites. The method is further
CC useful for predicting efficacy of a drug in treating viral infection,
CC plant genome shuffling, micropore manipulation, producing transgenic
CC animals, improvement of overexpressed genes for a desired phenotype and
CC reiterative pooling and breeding of higher organisms. The cells evolved
CC by the method are useful in molecular genetics. This sequence represents
CC an E. coli recA protein clone used in the method of the invention.

XX Sequence 358 AA;

Query Match 99.8%; Score 1674; DB 8; Length 358;
Best Local Similarity 99.7%; Pred. No. 7.7e-155;
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AIDENKQKALAAALQIEKQFGKSGIMRLGDRSMVDVTISTGSLSLDIALGAGGLPMGR 60
Db 7 AIDENKQKALAAALQIEKQFGKSGIMRLGDRSMVDVTISTGSLSLDIALGAGGLPMGR 66

Qy 61 IVEIYGPESSGKTTTLTQVIAAAQREGKTCFIDAEHALDPIYARKLGVDDINLLCSQPD 120
Db 67 IVEIYGPESSGKTTTLTQVIAAAQREGKTCFIDAEHALDPIYARKLGVDDINLLCSQPD 126

Qy 121 TGEQALEICDALARSGAVDVIWDSVAALTPKAEIEGIGSHMGLAARMMSQARMLAG 180
Db 127 TGEQALEICDALARSGAVDVIWDSVAALTPKAEIEGIGSHMGLAARMMSQARMLAG 186

Qy 181 NLKQSNLTLLIFINQIRMKIGVFNPGPPTTGGNALKFYASVRLDIRRIGAVKEGNNVGS 240
Db 187 NLKQSNLTLLIFINQIRMKIGVFNPGPPTTGGNALKFYASVRLDIRRIGAVKEGNNVGS 246

Qy 241 ETRVVKVKNKIAAPFKQAEFQILYEGINFGELVDLGVKEKLIBKAGAWYSYKGEKITGQ 300
Db 247 ETRVVKVKNKIAAPFKQAEFQILYEGINFGELVDLGVKEKLIBKAGAWYSYKGEKITGQ 306

Qy 301 GKANATAWLKONPETAKIEKKVRELLLSNPNSTP 335
Db 307 GKANATAWLKONPETAKIEKKVRELLLSNPNSTP 341

RESULT 8
ADQ87822

ID ADQ87822 standard; protein; 358 AA.

XX ADQ87822;

XX 09-SEP-2004 (first entry)

XX E. coli recA protein clone #1.

XX RecA; protoplast formation; gamete; heat tolerance; ethanol production;
KW ethanol tolerance; enzyme cofactor; NAD(P)H; glucose transport; meiosis;
KW heterokaryon; taxol; cyclosporin A; erythromycin; viral infection;
KW plant genome shuffling; micropore manipulation; reiterative pooling.

XX Escherichia coli.

XX AU2004200501-A1.

XX 04-MAR-2004.

XX 09-FEB-2004; 2004AU-00200501.

XX 09-FEB-2004; 2004AU-00200501.

XX (MAXY-) MAXYGEN INC.

PI Yuan L, Cox T, Bass S, Castle LA, Patten PA, Ness JB, Tobin M;
PI Affholter JA, Huisman G, Zhang Y, Kriebber CM, Subramanian V;
XX Minshull J, Stemmer WPC, Del Cardayre S;
DR WPI; 2004-507924/49.
XX N-PSDB; ADQ87815.

PT Evolving cells to acquire a desired property, by forming protoplasts of
PT different cells, fusing protoplasts to form hybrid protoplasts (HP),
PT producing regenerated cells, forming additional HP and producing
XX additional regenerated cells.

PS Disclosure; Fig 13; 196pp; English.

XX The invention relates to a method of evolving cells to acquire a desired
CC property by forming protoplasts of different cells, fusing the
CC protoplasts to form hybrid protoplasts, incubating the hybrid protoplasts
CC to produce regenerated cells, repeatedly forming protoplasts from
CC regenerated cells, fusing protoplasts to form hybrid protoplasts in which
CC genomes from protoplasts recombine to form additional hybrid genomes and
CC incubating additional hybrid protoplasts for producing additional
CC regenerated cells. The invention also relates to a method of producing a
CC library of diverse multicellular organisms involving providing a pool of
CC male gametes and a pool of female gametes, where one of the male pool or
CC the female pool comprises several different gametes derived from
CC different strains of a species or different species and the male gametes
CC fertilise the female gametes, permitting at least a portion of the
CC resulting fertilised gametes to grow into reproductively viable
CC organisms, repeatedly crossing the reproductively viable organisms to
CC produce a library of diverse organisms and selecting the library for a
CC desired trait or property. The methods are useful for evolving cells to
CC acquire desired properties such as heat tolerance, ethanol production,
CC ethanol tolerance, improved production and maintenance of enzyme
CC cofactors, improved production and maintenance of NAD(P)H and improved
CC glucose transport. The desired property is the expression of a protein, or
CC primary metabolite or secondary metabolite, the secretion of a protein or
CC secondary metabolite, capacity for meiosis or compatibility to form a
CC heterokaryon with another strain. The secondary metabolite is chosen from
CC taxol, cyclosporin A and erythromycin. The method is also useful for
CC acquisition of desired properties such as enhanced recombination, or
CC gene copy number or gene reductivity and capacity for expression and/or
CC secretion of proteins or secondary metabolites. The method is further
CC useful for predicting efficacy of a drug in treating viral infection,
CC plant genome shuffling, micropore manipulation, producing transgenic
CC animals, improvement of overexpressed genes for a desired phenotype and
CC reiterative pooling and breeding of higher organisms. The cells evolved
CC by the method are useful in molecular genetics. This sequence represents
CC an E. coli recA protein clone used in the method of the invention.

XX Sequence 358 AA;

Query Match 99.8%; Score 1674; DB 8; Length 358;

Best Local Similarity 99.7%; Pred. No. 7.7e-155;

Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AIDENKQKALAAALQIEKQFGKSGIMRLGDRSMVDVTISTGSLSLDIALGAGGLPMGR 60
Db 7 AIDENKQKALAAALQIEKQFGKSGIMRLGDRSMVDVTISTGSLSLDIALGAGGLPMGR 66

Qy 61 IVEIYGPESSGKTTTLTQVIAAAQREGKTCFIDAEHALDPIYARKLGVDDINLLCSQPD 120
Db 67 IVEIYGPESSGKTTTLTQVIAAAQREGKTCFIDAEHALDPIYARKLGVDDINLLCSQPD 126

Qy 121 TGEQALEICDALARSGAVDVIWDSVAALTPKAEIEGIGSHMGLAARMMSQARMLAG 180
Db 127 TGEQALEICDALARSGAVDVIWDSVAALTPKAEIEGIGSHMGLAARMMSQARMLAG 186

Qy 181 NLKQSNLTLLIFINQIRMKIGVFNPGPPTTGGNALKFYASVRLDIRRIGAVKEGNNVGS 240
Db 187 NLKQSNLTLLIFINQIRMKIGVFNPGPPTTGGNALKFYASVRLDIRRIGAVKEGNNVGS 246

Qy 241 ETRVVKVKNKIAAPFKQAEFQILYEGINFGELVDLGVKEKLIBKAGAWYSYKGEKITGQ 300

Db 247 ETRVVKVKNKIAAPFKQAEFOILYGEINFGVGLVDLGVKVKLEKAGAWSYSGEKIGQ 306

Qy 301 GKANATWLKNDPETAKEIEKKVRELLLSNPNSTP 335

Db 307 GKANATWLKNDPETAKEIEKKVRELLLSNPNSTP 341

RESULT 9

ADQ87821
ID ADQ87821 standard; protein; 358 AA.

XX

AC ADQ87821;

DT 09-SEP-2004 (first entry)

DE E. coli recA protein.

XX recA; protoplast formation; gamete; heat tolerance; ethanol production;
KW ethanol tolerance; enzyme cofactor; NAD(P)H; glucose transport; meiosis;
KW heterokaryon; taxol; cyclosporin A; erythromycin; viral infection;
KW plant genome shuffling; micropore manipulation; reiterative pooling.

OS Escherichia coli.

XX AU2004200501-A1.

XX 04-MAR-2004.

XX 09-FEB-2004; 2004AU-00200501.

XX 09-FEB-2004; 2004AU-00200501.

XX (MAXY-) MAXYGEN INC.

PI Yuan L, Cox T, Bass S, Castle LA, Patten PA, Ness JE, Tobin M;
FI Affholter JA, Huismann G, Zhang Y, Krebber CM, Subramanian V;
PI Minshull J, Stemmer WPC, Del Cardayre S;

XX WPI; 2004-507924/49.

XX N-PSDB; ADQ87814.

XX Evolving cells to acquire a desired property, by forming protoplasts of
PT different cells, fusing protoplasts to form hybrid protoplasts (HP),
PT producing regenerated cells, forming additional HP and producing
PT additional regenerated cells.

XX Disclosure; Fig 13; 196pp; English.

XX The invention relates to a method of evolving cells to acquire a desired
CC property by forming protoplasts of different cells, fusing the
CC protoplasts to form hybrid protoplasts, incubating the hybrid protoplasts
CC to produce regenerated cells, repeatedly forming protoplasts from
CC regenerated cells, fusing protoplasts to form hybrid protoplasts in which
CC genomes from protoplasts recombine to form additional hybrid genomes and
CC incubating additional hybrid protoplasts for producing additional
CC regenerated cells. The invention also relates to a method of producing a
CC library of diverse multicellular organisms involving providing a pool of
CC male gametes and a pool of female gametes, where one of the male pool or
CC the female pool comprises several different gametes derived from
CC different strains of a species or different species and the male gametes
CC fertilise the female gametes, permitting at least a portion of the
CC resulting fertilised gametes to grow into reproductively viable
CC organisms, repeatedly crossing the reproductively viable organisms to
CC produce a library of diverse organisms and selecting the library for a
CC desired trait or property. The methods are useful for evolving cells to
CC acquire desired properties such as heat tolerance, ethanol production,
CC ethanol tolerance, improved production and maintenance of enzyme
CC cofactors, improved production and maintenance of NAD(P)H and improved
CC glucose transport. The desired property is the expression of a protein,
CC primary metabolite or secondary metabolite, the secretion of a protein or
CC secondary metabolite, capacity for meiosis or compatibility to form a
CC heterokaryon with another strain. The secondary metabolite is chosen from
CC taxol, cyclosporin A and erythromycin. The method is also useful for

CC acquisition of desired properties such as enhanced recombinationity,
CC gene copy number or gene reductivity and capacity for expression and/or
CC secretion of proteins or secondary metabolites. The method is further
CC useful for predicting efficacy of a drug in treating viral infection,
CC plant genome shuffling, micropore manipulation, producing transgenic
CC animals, improvement of overexpressed genes for a desired phenotype and
CC reiterative pooling and breeding of higher organisms. The cells evolved
CC by the method are useful in molecular genetics. This sequence represents
CC the E. coli recA protein used in the method of the invention.

XX Sequence 358 AA;

Qy Query Match 99.8%; Score 1674; DB 8; Length 358;
Best Local Similarity 99.7%; Pred. No. 7.7e-155;
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AIDENKQKALAAALGOIEKQFGKGSIMRLGSDRSDVKTISTGSLSLDIALGAGGLPMGR 60

Db 7 AIDENKQKALAAALGOIEKQFGKGSIMRLGSDRSDVKTISTGSLSLDIALGAGGLPMGR 66

Qy 61 IVEIYGPESSGKTTLTLOVIAAAOREGKTCAPIDAEHALDPIYARKLGVDIDNLLCSQPD 120

Db 67 IVEIYGPESSGKTTLTLOVIAAAOREGKTCAPIDAEHALDPIYARKLGVDIDNLLCSQPD 126

Qy 121 TGEQALEICDALARSGAVDVIIVDSVAALTPKABIEGEIGDSHMLAARMMSQAMRKLKAG 180

Db 127 TGEQALEICDALARSGAVDVIIVDSVAALTPKABIEGEIGDSHMLAARMMSQAMRKLKAG 186

Qy 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGCGNALKFYASVRLDIRRIGAVKEGNNVGS 240

Db 187 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGCGNALKFYASVRLDIRRIGAVKEGNNVGS 246

Qy 241 ETRVVKVKNKIAAPFKQAEFOILYGEINFGVGLVDLGVKVKLEKAGAWSYSGEKIGQ 300

Db 247 ETRVVKVKNKIAAPFKQAEFOILYGEINFGVGLVDLGVKVKLEKAGAWSYSGEKIGQ 306

Qy 301 GKANATWLKNDPETAKEIEKKVRELLLSNPNSTP 335

Db 307 GKANATWLKNDPETAKEIEKKVRELLLSNPNSTP 341

RESULT 10

AY68831

ID AY68831 standard; protein; 358 AA.

XX

AC AY68831;

DT 16-MAY-2000 (first entry)

DE Amino acid sequence of a hyperrecombinogenic recA protein clone 6.

XX recA; hyperrecombinogenic variant; male gamete; female gamete;

XX heat tolerance; ethanol production; ethanol tolerance; taxol;

XX cyclosporin A; erythromycin; meiosis.

XX Synthetic.

XX Escherichia coli.

XX Key Location/Qualifiers

XX Misc-difference 156

XX /note= "encoded by TCG"

XX WO200004190-A1.

XX 27-JAN-2000.

XX 15-JUL-1999; 99WO-US015972.

XX 15-JUL-1998; 98US-00116188.

XX (MAXY-) MAXYGEN INC.

XX Del Cardayre S, Tobin M, Stemmer WPC, Ness JE, Minshull J;

PI Patten PA, Subramanian V, Castle LA, Krebber CM, Bass S, Zhang Y;
 PI Cox T, Huisman G, Yuan L, Affholter JA;
 DR WPI; 2000-182446/16.
 DR N-PSDB; AAZ60613.
 XX Evolution of whole cells and organism by iterative cycles of
 PT recombination and selection and screening for acquisition of desired
 PT properties.
 XX Example 1; Fig 13; 197pp; English.
 PS The present sequence represents a hyperrecombinogenic recA protein. The
 CC wild type recA protein (AAV68827) was used to produce hyperrecombinogenic
 CC variants, using the method of the invention. The specification describes
 CC a method for producing a library of diverse multicellular organisms using
 CC pools of male and female gametes. At least one of the male pool or female
 CC pools comprises a number of different gametes derived from different
 CC strains of a species or of a different species. The viable organisms
 CC produced from the fertilized gametes are repeatedly crossed to produce a
 CC library of diverse organisms, which are selected for a desired trait or
 CC property. The methods can be used to evolve cells to acquire a desired
 CC property such as heat tolerance, ethanol production or tolerance, acid,
 CC improved production and maintenance of enzyme cofactors or NAD(P)H and
 CC improved glucose transport. The desired property may be expression of a
 CC protein or primary or secondary metabolite. Alternatively the desired
 CC property is secretion of a protein or secondary metabolite, chosen from
 CC taxol, cyclosporin A and erythromycin. The desired property may be a
 CC capacity for meiosis or compatibility to form a heterokaryon with another
 CC strain
 XX Sequence 358 AA;

Query Match 99.7%; Score 1673; DB 3; Length 358;
 Best Local Similarity 99.4%; Pred. No. 9.7e-155;
 Matches 333; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMDVKTISTGSLSLDIALGAGGLPMGR 60
 DB 7 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMDVKTISTGSLSLDIALGAGGLPMGR 66
 QY 61 IVEIYGPESSGKTTTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
 DB 67 IVEIYGPESSGKTTTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 126
 QY 121 TGEQALEICDALARSGAVDVIWDSVAALTPKAEIEGIGDSHMGSLAARMMSQAMRKLKAG 180
 DB 127 TGEQALEICDALARSGAVDVIWDSVAALTPKAEIEGIGDSHMGSLAARMMSQAMRKLKAG 186
 QY 181 NLKQNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
 DB 187 NLKQNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 246
 QY 241 ETRVKVKNKTAAPFKQAEFQILYEGEINFYGVLDVGVKEKLEKAGAWTSYKGEKIGQ 300
 DB 247 ETRVKVKNKTAAPFKQAEFQILYEGEINFYGVLDVGVKEKLEKAGAWTSYKGEKIGQ 306
 QY 301 GKANATAWLKONPETAKEIEKKVRELLLSNPNSTP 335
 DB 307 GKANATAWLKONPETAKEIEKKVRELLLSNPNSTP 341

RESULT 11
 AAW64213
 ID AAW64213 standard; protein; 358 AA.
 XX AAW64213;
 XX

28-APR-1999 (first entry)
 DE New minshall recA protein.
 XX

Escherichia coli; recA protein; hyperrecombinogenic variant; Minshall;

KW recursive sequence recombination; evolution.
 XX Escherichia coli.
 OS WO9831837-A1.
 PN 23-JUL-1998.
 PD 16-JAN-1998; 98WO-US000852.
 PF 17-JAN-1997; 97US-0035054P.
 PR (MAXY-) MAXYGEN INC.
 PA Delcardayre SB, Tobin MB, Stemmer WPC, Ness JE, Minshull J;
 PI Patten P;
 XX WPI; 1998-427565/60.
 DR N-PSDB; AAV44285.
 XX Evolution of whole cells and organisms by recursive DNA sequence
 PT recombination in cells to evolve cells having acquired desired function,
 PT useful in methods for predicting the efficacy of a drug in treating viral
 PT or pathogenic infections.
 XX Example 1; Fig 13; 125pp; English.
 PS The present invention provides methods employing iterative cycles of
 CC recombination and selection/screening for evolution of whole cells and
 CC organisms toward acquisition of desired properties, e.g. enhanced
 CC recombogenicity, genome copy number, and capacity for expression and/or
 CC secretion of proteins and secondary metabolites. The present sequence
 CC represents a wild-type recA protein (designated new Minshall), from an
 CC example of the present invention
 XX Sequence 358 AA;

Query Match 99.5%; Score 1670; DB 2; Length 358;
 Best Local Similarity 99.4%; Pred. No. 1.9e-154;
 Matches 333; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMDVKTISTGSLSLDIALGAGGLPMGR 60
 DB 7 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMDVKTISTGSLSLDIALGAGGLPMGR 66
 QY 61 IVEIYGPESSGKTTTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
 DB 67 IVEIYGPESSGKTTTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 126
 QY 121 TGEQALEICDALARSGAVDVIWDSVAALTPKAEIEGIGDSHMGSLAARMMSQAMRKLKAG 180
 DB 127 TGEQALEICDALARSGAVDVIWDSVAALTPKAEIEGIGDSHMGSLAARMMSQAMRKLKAG 186
 QY 181 NLKQNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
 DB 187 NLKQNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 246
 QY 241 ETRVKVKNKTAAPFKQAEFQILYEGEINFYGVLDVGVKEKLEKAGAWTSYKGEKIGQ 300
 DB 247 ETRVKVKNKTAAPFKQAEFQILYEGEINFYGVLDVGVKEKLEKAGAWTSYKGEKIGQ 306
 QY 301 GKANATAWLKONPETAKEIEKKVRELLLSNPNSTP 335
 DB 307 GKANATAWLKONPETAKEIEKKVRELLLSNPNSTP 341

RESULT 12
 AAY68828
 ID AAY68828 standard; protein; 358 AA.
 XX AAY68828;
 XX
 DT 16-MAY-2000 (first entry)

XX	Amino acid sequence of a hyperrecombinogenic recA protein clone 2.
DE	
XX	recA; hyperrecombinogenic variant; male gamete; female gamete;
KW	heat tolerance; ethanol production; ethanol tolerance; metabolite; taxol;
KW	cyclosporin A; erythromycin; meiosis.
XX	Synthetic.
OS	Escherichia coli.
XX	WO200004190-A1.
PN	
XX	27-JAN-2000.
PD	
XX	15-JUL-1999; 99WO-US015972.
Pf	
XX	15-JUL-1998; 98US-00116188.
PR	
XX	(MAXY-) MAXYGEN INC.
PA	
XX	Del Cardayre S, Tobin M, Stemmer WPC, Ness JE, Minshull J;
PI	Patten PA, Subramanian V, Castle LA, Krebber CM, Bass S, Zhang Y;
PI	Cox T, Huisman G, Yuan L, Affholter JA;
XX	WPI; 2000-182446/16.
DR	N-PSDE; AAZ60610.
DR	
XX	Evolution of whole cells and organism by iterative cycles of
PT	recombination and selection and screening for acquisition of desired
PT	properties.
XX	Example 1; Fig 13; 197pp; English.
XX	The present sequence represents a hyperrecombinogenic recA protein. The
CC	wild type recA protein (AAV68927) was used to produce hyperrecombinogenic
CC	variants, using the method of the invention. The specification describes
CC	a method for producing a library of diverse multicellular organisms using
CC	pools of male and female gametes. At least one of the male pool or female
CC	pools comprises a number of different gametes derived from different
CC	strains of a species or of a different species. The viable organisms
CC	produced from the fertilized gametes are repeatedly crossed to produce a
CC	library of diverse organisms, which are selected for a desired trait or
CC	property. The methods can be used to evolve cells to acquire a desired
CC	property such as heat tolerance, ethanol production or tolerance, acid,
CC	improved production and maintenance of enzyme cofactors or NAD(P)H and
CC	improved glucose transport. The desired property may be expression of a
CC	protein or primary or secondary metabolite. Alternatively the desired
CC	property is secretion of a protein or secondary metabolite, chosen from
CC	taxol, cyclosporin A and erythromycin. The desired property may be a
CC	capacity for meiosis or compatibility to form a heterokaryon with another
CC	strain
XX	
SQ	Sequence 358 AA;
	Query Match 99.5%; Score 1669; DB 3; Length 358;
	Best Local Similarity 99.1%; Pred. No. 2.4e-154;
	Matches 332; Conservative 2; Mismatches 1; Indels 0; Gaps 0
Qy	1 AIDENKOKALAAALGQIEKQFGKGSIMRLGEDRSDMDVTITSTGSLSDLTALGAGGLPMGR 60
Dd	7 AIDENKOKALTALGQIEKQFGKGSIMRLGEDRSDMDVTITSTGSLSDLTALGAGGLPMGR 66
Qy	61 IVEIYGPESSCKTTLTLOVIAAQQREBKTCFIDAETHALDPTIYARKLGVDINDTLCSQPD 120
Dd	67 IVEIYGPESSCKTTLTUQVIAAQQREBKTCFIDAETHALDPTIYARKLGVDINDTLCSQPD 126
Qy	121 TGEQALEICDALARGAVDVIVVDSVAALTPKAEITEGEBIGSHMGLAARWMSQAMRKLAG 180
Dd	127 TGEQALEICDALARGAVDVIVVDSVAALTPKAEITEGEBIGSHMGLAARWMSQAMRKLAG 186
Qy	181 NLKQSNTLLIFINQIRMKIGVMFNPPETTTGGNALKFYASVRLDIRRIQAVKEGVNVVGS 240
Dd	187 NLKQSNTLLIFINQIRMKIGVMFNPPETTTGGNALKFYASVRLDIRRIQAVKEGVNVVGS 246

CC secondary metabolite, capacity for meiosis or compatibility to form a
CC heterokaryon with another strain. The secondary metabolite is chosen from
CC taxol, cyclosporin A and erythromycin. The method is also useful for
CC acquisition of desired properties such as enhanced recombination, gene
CC secretion of proteins or secondary metabolites. The method is further
CC useful for predicting efficacy of a drug in treating viral infection,
CC plant genome shuffling, microprobe manipulation, producing transgenic
CC animals, improvement of overexpressed genes for a desired phenotype and
CC reiterative pooling and breeding of higher organisms. The cells evolved
CC by the method are useful in molecular genetics. This sequence represents
CC an E. coli recA protein clone used in the method of the invention.
XX
SQ

Query Match 99.5%; Score 1669; DB 8; Length 358;
Best Local Similarity 99.1%; Pred. No. 2.4e-154;
Matches 332; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSDVKTISTGSLDLIALGAGGLPMGR 60
DB 7 AIDENKQKALATAGQIEKQFGKGSIMRLGEDRSDVKTISTGSLDLIALGAGGLPMGR 66
QY 61 IVEIYGPSSGKTTTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVDDNLLCSQPD 120
DB 67 IVEIYGPSSGKTTTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVDDNLLCSQPD 126
QY 121 TGEQALEICDALARSGAVDVIVDSVAALTTPKAEIEGIGDSHMGLAARMMSQAMRKLAG 180
DB 127 TGEQALEICDALARSGAVDVIVDSVAALTTPKAEIEGIGDSHMGLAARMMSQAMRKLAG 186
QY 181 NLKQSNLTLLIFINQIRMKIGWFGNPNPTTTCGNALKFYASVRLDIRRIGAVKEGENVYGS 240
DB 187 NLKQSNLTLLIFINQIRMKIGWFGNPNPTTTCGNALKFYASVRLDIRRIGAVKEGENVYGS 246
QY 241 ETRVVKVNKNTAAPFKQAEFOILYEGEINFYGVLDLGVKEKLIBKAGAWTSYKGEKIGQ 300
DB 247 ETRVVKVNKNTAAPFKQAEFOILYEGEINFYGVLDLGVKEKLIBKAGAWTSYKGEKIGQ 306
QY 301 GKANATAMLKONPETAKIEKKVRELLLSNPNSTP 335
DB 307 GKANATAMLKONPETAKIEKKVRELLLSNPNSTP 341

RESULT 14
ADQ87826
ID ADQ87826 standard; protein; 358 AA.
XX
AC ADQ87826;
XX
XX
DT 09-SEP-2004 (first entry)
XX
XX
DE E. coli recA protein clone #5.
XX
KW RecA; protoplast formation; gamete; heat tolerance; ethanol production;
KW ethanol tolerance; enzyme cofactor; NAD(P)H; Glucose transport; meiosis;
KW heterokaryon; taxol; cyclosporin A; erythromycin; viral infection;
KW plant genome shuffling; microprobe manipulation; reiterative pooling.
XX
OS Escherichia coli.
XX
FN AU2004200501-A1.
XX
XX
PD 04-MAR-2004.
XX
XX
PF 09-FEB-2004; 2004AU-00200501.
XX
PR 09-FEB-2004; 2004AU-00200501.
XX
PA (MAXY-) MAXYGEN INC.
XX
PI Yuan L, Cox T, Bass S, Castle LA, Patten PA, Ness JE, Tobin M;
PI Affholter JA, Huisman G, Zhang Y, Krebber CM, Subramanian V;

PI Minshull J, Stemmer WPC, Del Cardayre S;
XX
XX WPI; 2004-507924/49.
DR N-PSDB; ADQ87819.
XX
XX
PT Evolving cells to acquire a desired property, by forming protoplasts of
PT different cells, fusing protoplasts to form hybrid protoplasts (HP),
PT producing regenerated cells, forming additional HP and producing
PT additional regenerated cells.
XX
XX
FS Disclosure; Fig 13; 196pp; English.
XX
XX

The invention relates to a method of evolving cells to acquire a desired
property by forming protoplasts of different cells, fusing the
protoplasts to form hybrid protoplasts, incubating the hybrid protoplasts
to produce regenerated cells, repeatedly forming protoplasts from
regenerated cells, fusing protoplasts to form hybrid protoplasts in which
genomes from protoplasts recombine to form additional hybrid genomes and
incubating additional hybrid protoplasts for producing additional
regenerated cells. The invention also relates to a method of producing a
library of diverse multicellular organisms involving providing a pool of
male gametes and a pool of female gametes, where one of the male pool or
the female pool comprises several different gametes derived from
different strains of a species or different species and the male gametes
fertilise the female gametes, permitting at least a portion of the
resulting fertilised gametes to grow into reproductively viable
organisms, repeatedly crossing the reproductively viable organisms to
produce a library of diverse organisms and selecting the library for a
desired trait or properties. The methods are useful for evolving cells to
acquire desired properties such as heat tolerance, ethanol production,
ethanol tolerance, improved production and maintenance of enzyme
cofactors, improved production and maintenance of NAD(P)H and improved
glucose transport. The desired property is the expression of a protein,
primary metabolite or secondary metabolite, the secretion of a protein or
heterokaryon with another strain. The secondary metabolite is chosen from
taxol, cyclosporin A and erythromycin. The method is also useful for
acquisition of desired properties such as enhanced recombination, gene
copy number or gene reductivity and capacity for expression and/or
secretion of proteins or secondary metabolites. The method is further
useful for predicting efficacy of a drug in treating viral infection,
plant genome shuffling, microprobe manipulation, producing transgenic
animals, improvement of overexpressed genes for a desired phenotype and
reiterative pooling and breeding of higher organisms. The cells evolved
by the method are useful in molecular genetics. This sequence represents
an E. coli recA protein clone used in the method of the invention.

Sequence 358 AA;

Query Match 99.2%; Score 1665; DB 8; Length 358;
Best Local Similarity 99.1%; Pred. No. 5.9e-154;
Matches 332; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSDVKTISTGSLDLIALGAGGLPMGR 60
DB 7 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSDVKTISTGSLDLIALGAGGLPMGR 66
QY 61 IVEIYGPSSGKTTTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVDDNLLCSQPD 120
DB 67 IVEIYGPSSGKTTTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVDDNLLCSQPD 126
QY 121 TGEQALEICDALARSGAVDVIVDSVAALTTPKAEIEGIGDSHMGLAARMMSQAMRKLAG 180
DB 127 TGEQALEICDALARSGAVDVIVDSVAALTTPKAEIEGIGDSHMGLAARMMSQAMRKLAG 186
QY 181 NLKQSNLTLLIFINQIRMKIGWFGNPNPTTTCGNALKFYASVRLDIRRIGAVKEGENVYGS 240
DB 187 NLKQSNLTLLIFINQIRMKIGWFGNPNPTTTCGNALKFYASVRLDIRRIGAVKEGENVYGS 246
QY 241 ETRVVKVNKNTAAPFKQAEFOILYEGEINFYGVLDLGVKEKLIBKAGAWTSYKGEKIGQ 300
DB 247 ETRVVKVNKNTAAPFKQAEFOILYEGEINFYGVLDLGVKEKLIBKAGAWTSYKGEKIGQ 306

QY 301 GKANATAWLKONPETAKEIEKKVRELLLSNPNSTP 335
DB 307 GKANATAWLKONPETAKEIEKKVRELLLSNPNSTP 341

RESULT 15

AAW64215
ID AAW64215 standard; protein; 358 AA.

XX AC AAW64215;

XX XX 28-APR-1999 (first entry)

XX DE Hyperrecombinogenic variant recA protein clone 4.

XX KW Escherichia coli; recA protein; hyperrecombinogenic variant; Minshall;
XX OS recursive sequence recombination; evolution.

XX OS Escherichia coli.
XX OS Synthetic.

XX PN WO9831837-A1.

XX PD 23-JUL-1998.

XX XX 16-JAN-1998; 98WO-US000852.

XX PF 17-JAN-1997; 97US-0035054P.

XX PR (MAXY-) MAXYGEN INC.

XX PA Delcardayre SB, Tobin MB, Stemmer WPC, Ness JE, Minshall J;
XX PI Patten P;

XX DR WPI; 1998-427565/60.
XX DR N-FSDB; AAW44287.

XX XX Evolution of whole cells and organisms by recursive DNA sequence
PT recombination in cells to evolve cells having acquired desired function,
PT useful in methods for predicting the efficacy of a drug in treating viral
PT or pathogenic infections.

XX PS Example 1; Fig 13; 125pp; English.

XX CC The present invention provides methods employing iterative cycles of
CC recombination and selection/screening for evolution of whole cells and
CC organisms toward acquisition of desired properties, e.g. enhanced
CC recombinogenicity, genome copy number, and capacity for expression and/or
CC secretion of proteins and secondary metabolites. The present sequence
CC represents a hyperrecombinogenic variant of a recA protein (see
CC AAW64213), from an example of the present invention

XX SQ Sequence 358 AA;

Query Match 99.2%; Score 1664; DB 2; Length 358;
Best Local Similarity 98.8%; Pred. No. 7.4e-154;
Matches 331; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 AIDENKQKALAAALGOIEKQFGKGSIMRLGEDRSDMVKTISTGSLSLDIALGAGGLPMGR 60

DB 7 AIDENKQKALAAALGOIEKQFGKGSIMRLGEDRSDMVKTISTGSLSLDIALGAGGLPMGR 66

QY 61 IVEIYGPESSCKTTLTQVIAAAREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 120

DB 67 IVEIYGPESSCKTTLTQVIAAAREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 126

QY 121 TGEQALEICDALARSGAVDVIVVDSVAALTPKAEIEGIDSHMGLAARMMSQAMRKLKAG 180

DB 127 TGEQALEICDALARSGAVDVIVVDSVAALTTPKASIEGIDSHMGLAARMMSQAMRKLKAG 186

QY 181 NLKQSNLTLLIPINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240

DB 187 NLKQSNLTLLIPINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 246

QY 241 ETRVKKVVKNKIAAPFKQAEFQILYGEGINFYGELVDLGVKEKLI EKAGAWSYKGEKIGQ 300
DB 247 ETRVKKVVKNKIAAPFKQAEFQILYGEGINFYGELVDLGVKEKLI EKAGAWSYKGEKVGQ 306

QY 301 GKANATAWLKONPETAKEIEKKVRELLLSNPNSTP 335

DB 307 GKANATAWLKONPETAKEIEKKVRELLLSNPNSTP 341

Search completed: February 16, 2006, 01:18:52
Job time : 133 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: February 16, 2006, 01:19:09 ; Search time 25.5 Seconds
(without alignments)
1264.024 Million cell updates/sec

Title: US-10-733-782-3
Perfect score: 1678
Sequence: 1 AIDENKOKALAAALGQIEKQ.....AKIEKVKRELLSNPNSTP 335
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1674	99.8	353	1	ROECA
2	1674	99.8	353	2	Recombination prot
3	1662	99.0	353	2	Reca protein [impo
4	1637	97.6	353	2	hypothetical prote
5	1530	91.2	356	2	Reca protein [impo
6	1528	91.1	354	2	Recombination prot
7	1526	90.9	356	2	recombination prot
8	1500	89.4	355	1	recombination prot
9	1432	85.3	412	2	Reca protein VC054
10	1413	84.2	354	2	recombination prot
11	1408.5	83.9	348	2	recombination prot
12	1328	79.1	354	2	recombination prot
13	1319	78.6	353	2	recombination prot
14	1268	75.6	346	1	Reca protein PA361
15	1265	75.4	349	2	recombination prot
16	1263.5	75.3	344	2	recombination prot
17	1255	74.8	355	2	Reca protein - Pse
18	1253	74.7	349	2	recombination prot
19	1250	74.5	352	2	recombination prot
20	1246	74.3	352	2	recombination prot
21	1234.5	73.6	348	1	Recombination prot
22	1226.5	73.1	342	2	recombination prot
23	1213.5	72.3	348	2	Reca protein NMA14
24	1212.5	72.3	348	2	Reca protein NMA16
25	1209	72.1	378	2	Reca protein [impo
26	1200.5	71.5	347	1	recombination prot
27	1199.5	71.5	348	2	recombination prot
28	1199.5	71.5	351	2	recombination prot
29	1196.5	71.3	363	2	recombination prot

ALIGNMENTS

RESULT 1

ROECA

recombination protein reca [validated] - Escherichia coli (strain K-12)

N;Alternate names: recombinase A

C;Species: Escherichia coli

C;Date: 31-Jul-1980 #sequence revision 14-Nov-1997 #text change 09-Jul-2004

C;Accession: G65049; A93847; A93846; S11931; S63525; S69139; S63979; A03548

R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;

.A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A;Title: The complete genome sequence of Escherichia coli K-12.

A;Reference number: A64720; MUID:97426617; PMID:9278503

A;Accession: G65049

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-353 <BLAT>

A;Cross-references: UNIPROT:P03017; UNIPARC:UPI0000112C1C; GB:AE000354; GB:U000096; NID

A;Experimental source: strain K-12, substrain MG1655

R;Horii, T.; Ogawa, T.; Ogawa, H.

Proc. Natl. Acad. Sci. U.S.A. 77, 313-317, 1980

A;Title: Organization of the reca gene of Escherichia coli.

A;Reference number: A93847; MUID:80145618; PMID:6244554

A;Accession: A93847

A;Molecule type: DNA

A;Residues: 2-353 <HOR>

A;Cross-references: UNIPARC:UPI0000112C17; GB:V00328; NID:942672; PIDN:CAA23618.1; PID

R;Sancar, A.; Stacheliek, C.; Konigsberg, W.; Rupp, W.D.

Proc. Natl. Acad. Sci. U.S.A. 77, 2611-2615, 1980

A;Title: Sequences of the reca gene and protein.

A;Reference number: A93846; MUID:80234673; PMID:6930655

A;Accession: A93846

A;Molecule type: DNA

A;Residues: 2-353 <SAN>

A;Cross-references: UNIPARC:UPI0000112C17; GB:V00328; NID:942672; PIDN:CAA23618.1; PIDN:CAA23

R;Zhao, X.J.; McEntee, K.

Mol. Gen. Genet. 222, 369-376, 1990

A;Title: DNA sequence analysis of the reca genes from Proteus vulgaris, Erwinia carotov

A;Reference number: S11931; MUID:91109725; PMID:2274037

A;Accession: S11931

A;Molecule type: DNA

A;Residues: 2-353 <ZHA>

A;Cross-references: UNIPARC:UPI0000112C17; GB:X55552; NID:942678

A;Note: this ORF is not annotated in GenBank entry ECRECAGEN, release 109.0

R;Morimatsu, K.; Horii, T.

Eur. J. Biochem. 234, 695-705, 1995

A;Title: DNA-binding surface of RecA protein. Photochemical cross-linking of the first

A;Reference number: S63525; MUID:96096752; PMID:8529655

A;Accession: S63525

A;Status: preliminary

A;Molecule type: protein

A;Residues: 65-69;90-97;179-184;200-207;258-265;304-311;323-331 <MORI>

A;Cross-references: UNIPARC:UPI0001748A7; UNIPARC:UPI00001748A6; UNIPARC:UPI00001748A5

R.Morimatsu, K.; Hori, T.
Eur. J. Biochem. 228, 772-778, 1995
A:Title: The DNA-binding site of the RecA protein. Photochemical cross-linking of Tyr103
A:Reference number: S69129; MUID:95255284; PMID:7737176
A:Accession: S69129
A:Status: preliminary
A:Molecule type: protein
A:Residues: 90-108,180-184 <MOR2>
A:CROSS-references: UNIPARC:UPI00001748AE
R:Gardner, R.V.; Voloshin, O.N.; Camerini-Otero, R.D.
Eur. J. Biochem. 233, 419-425, 1995
A:Title: The identification of the single-stranded DNA-binding domain of the Escherichia
A:Reference number: S63979; MUID:96067680; PMID:7588783
A:Accession: S63979
A:Molecule type: protein
A:Residues: 'XX',187-190,192-194 <GAR>
A:CROSS-references: UNIPARC:UPI00001748AF
R:Yu, X.; Egelman, E.H.
submitted to the Brookhaven Protein Data Bank, December 1996
A:Reference number: A67277; PDB:2REC
A:Contents: annotation; X-ray crystallography, 2.3 angstroms, residues 4-157;166-195;211
R:Aihara, H.; Ito, Y.; Kurumizaka, H.; Terada, T.; Yokoyama, S.; Shibata, T.
submitted to the Brookhaven Protein Data Bank, January 1997
A:Reference number: A67455; PDB:1AA3
A:Contents: annotation; conformation by (1)H- and (15)N-NMR, residues 269-331
C:Genetics:
A:Gene: recA
A:Map position: 58 min
C:Function:
A:Description: plays an essential role in homologous recombination, in induction of the
C:Superfamily: recombination protein recA
C:Keywords: ATP; DNA binding; DNA recombination; DNA repair; nucleotide binding; P-loop;
F:67-74/Region: nucleotide-binding motif A (P-loop)
F:141-146/Region: nucleotide-binding motif B
F:73/Binding site: ATP (Lys) #status predicted

Query Match 99.8%; Score 1674; DB 1; Length 353;
Best Local Similarity 99.7%; Pred. No. 3.6e-106;
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVKTISTGSLSLDIALGAGGLPMGR 60
Db 2 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVKTISTGSLSLDIALGAGGLPMGR 61
Qy 61 IVEIYGPSSGKTTLTQVIAAAQREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 62 IVEIYGPSSGKTTLTQVIAAAQREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 121
Qy 121 TGEQALEICDALARSGAVDIVVDSVAALTPKAEIEGIGDSHMGLAARMMSQAMRKLKAG 180
Db 122 TGEQALEICDALARSGAVDIVVDSVAALTPKAEIEGIGDSHMGLAARMMSQAMRKLKAG 181
Qy 181 NLKQNTLLIFINQIRMKIGVFNPGPETTTGKNALKFYASVRLDIRRIGAVKEGENVVGS 240
Db 182 NLKQNTLLIFINQIRMKIGVFNPGPETTTGKNALKFYASVRLDIRRIGAVKEGENVVGS 241
Qy 241 ETRVKVKNKIAAPFKQAEFQILYGEINFGELVDLGKVKLEIKAGAWTSYKGEKIGQ 300
Db 242 ETRVKVKNKIAAPFKQAEFQILYGEINFGELVDLGKVKLEIKAGAWTSYKGEKIGQ 301
Qy 301 GKANATAWLKONPETAKIEKKVRELLLSNPSTP 335
Db 302 GKANATAWLKONPETAKIEKKVRELLLSNPSTP 336

RESULT 2
D91073
RecA protein [imported] - Escherichia coli (strain O157:H7, substrain RMD 0509952)
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: D91073
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: D91073
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-353 <HAY>
A:CROSS-references: UNIPROT:P03017; UNIPARC:UPI0000112C1C; GB:BA0000007; PIDN:BA036979.1;
A:Experimental source: strain O157:H7, substrain RMD 0509952
C:Genetics:
A:Gene: EC63556
C:Superfamily: recombination protein recA

Query Match 99.8%; Score 1674; DB 2; Length 353;
Best Local Similarity 99.7%; Pred. No. 3.6e-106;
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVKTISTGSLSLDIALGAGGLPMGR 60
Db 2 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVKTISTGSLSLDIALGAGGLPMGR 61
Qy 61 IVEIYGPSSGKTTLTQVIAAAQREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 62 IVEIYGPSSGKTTLTQVIAAAQREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 121
Qy 121 TGEQALEICDALARSGAVDIVVDSVAALTPKAEIEGIGDSHMGLAARMMSQAMRKLKAG 180
Db 122 TGEQALEICDALARSGAVDIVVDSVAALTPKAEIEGIGDSHMGLAARMMSQAMRKLKAG 181
Qy 181 NLKQNTLLIFINQIRMKIGVFNPGPETTTGKNALKFYASVRLDIRRIGAVKEGENVVGS 240
Db 182 NLKQNTLLIFINQIRMKIGVFNPGPETTTGKNALKFYASVRLDIRRIGAVKEGENVVGS 241
Qy 241 ETRVKVKNKIAAPFKQAEFQILYGEINFGELVDLGKVKLEIKAGAWTSYKGEKIGQ 300
Db 242 ETRVKVKNKIAAPFKQAEFQILYGEINFGELVDLGKVKLEIKAGAWTSYKGEKIGQ 301
Qy 301 GKANATAWLKONPETAKIEKKVRELLLSNPSTP 335
Db 302 GKANATAWLKONPETAKIEKKVRELLLSNPSTP 336

RESULT 3
H85917
hypochemical protein recA [imported] - Escherichia coli (strain O157:H7, substrain EDL93
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: H85917
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glaesner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamouais, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: H85917
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-353 <STO>
A:CROSS-references: UNIPARC:UPI00001658E3; GB:A8005174; NID:gl2517139; PIDN:AAG57804.1;
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: recA
C:Superfamily: recombination protein recA

Query Match 99.0%; Score 1662; DB 2; Length 353;
Best Local Similarity 99.1%; Pred. No. 2.3e-105;
Matches 332; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVKTISTGSLSLDIALGAGGLPMGR 60
Db 2 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVKTISTGSLSLDIALGAGGLPMGR 61
Qy 61 IVEIYGPSSGKTTLTQVIAAAQREGKTCFIDAEHALDPIYARKLGVDIDNLLCSQPD 120

Db 62 IVEIYGPESSGKTTLTQVIAAAREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 121

Qy 121 TGEQALEICDALARSGAVDVIVVDSVAALTPKAEIEGIDGSHMGLAARMMWSQAMRKLKAG 180

Db 122 TGEQALEICDALARSGAVDVIVVDSVAALTPKAEIEGIDGSHMGLAARMMWSQAMRKLKAG 181

Qy 181 NLKQSNLTLLIFINQIRMKIGVMFGNPTTTGGNALKFYASVRLDIRRIGAVKEGNNVVG 240

Db 182 NLKQSNLTLLIFINQIRMKIGVMFGNPTTTGGNALKFYASVRLDIRRIGAVKEGNNVVG 241

Qy 241 ETRVKVVKNTAAAPFKQAEFOILYEGEINFGVGLVDLGVKEKLEKAGAWYSYKGEKIGQ 300

Db 242 ETRVKVVKNTAAAPFKQAEFOILYEGEINFGVGLVDLGVKEKLEKAGAWYSYKGEKIGQ 301

Qy 301 GKANATAWLKNDPETAKEIEKKVRELLLSNPNSTP 335

Db 302 GKANATAWLKNDPETAKEIEKKVRELLLSNPNSTP 336

RESULT 4

AH0843

ReCA protein [imported] - Salmonella enterica subsp. enterica serovar Typhi (strain CT18

C;Species: Salmonella enterica subsp. enterica serovar Typhi

A;Note: this species has also been called Salmonella typhi

C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002

C;Accession: AH0843

R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moulé, S.; O'Gaora, P.

Nature 413, 848-852, 2001

A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.

A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov

A;Reference number: AB0502; MUID:2153494; PMID:11677608

A;Accession: AH0843

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-353 <PAR>

A;Cross-references: UNIPARC:UPI000005A334; GB:AL513382; PIDN:CAD05935.1; PID:g16503906;

C;Genetics:

A;Gene: STV2950

C;Superfamily: recombination protein recA

Query Match 97.6%; Score 1637; DB 2; Length 353;

Best Local Similarity 97.0%; Pred. No. 1.2e-103;

Matches 325; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

ReCA protein [imported] - Yersinia pestis (strain CO92)

C;Species: Yersinia pestis

C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004

C;Accession: AG0401

R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.F deno-Tarrega, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; ll, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, Nature 413, 523-527, 2001

A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A;Reference number: AB0001; MUID:21470413; PMID:11586360

A;Accession: AG0401

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-356 <KUR>

A;Cross-references: UNIPROT:P37850; UNIPARC:UPI0000165AF8; GB:AL590842; PIDN:CAC92539.1

C;Genetics:

A;Gene: recA

C;Superfamily: recombination protein recA

Query Match 91.2%; Score 1530; DB 2; Length 356;

Best Local Similarity 91.6%; Pred. No. 2.1e-96;

Matches 304; Conservative 15; Mismatches 13; Indels 0; Gaps 0;

Qy 1 AIDENKOKALAAALGQIEKQFGKGSIMRLGEDRSMVKTISTGSLSDIALGAGGLPMGR 60

Db 2 AIDENKOKALAAALGQIEKQFGKGSIMRLGEDRSMVKTISTGSLSDIALGAGGLPMGR 61

Qy 61 IVEIYGPESSGKTTLTQVIAAAREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 120

Db 62 IVEIYGPESSGKTTLTQVIAAAREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 121

Qy 121 TGEQALEICDALARSGAVDVIVVDSVAALTPKAEIEGIDGSHMGLAARMMWSQAMRKLKAG 180

Db 122 TGEQALEICDALARSGAVDVIVVDSVAALTPKAEIEGIDGSHMGLAARMMWSQAMRKLKAG 181

Qy 181 NLKQSNLTLLIFINQIRMKIGVMFGNPTTTGGNALKFYASVRLDIRRIGAVKEGNNVVG 240

Db 182 NLKQSNLTLLIFINQIRMKIGVMFGNPTTTGGNALKFYASVRLDIRRIGAVKEGNNVVG 241

Qy 241 ETRVKVVKNTAAAPFKQAEFOILYEGEINFGVGLVDLGVKEKLEKAGAWYSYKGEKIGQ 300

Db 242 ETRVKVVKNTAAAPFKQAEFOILYEGEINFGVGLVDLGVKEKLEKAGAWYSYKGEKIGQ 301

Qy 301 GKANATAWLKNDPETAKEIEKKVRELLLSNPN 332

Db 302 GKANATYLNKENPALAAELDKLREMLNGN 333

RESULT 6

S31481

recombination protein recA - Enterobacter agglomerans

N;Alternate names: recombinase A

C;Species: Enterobacter agglomerans

C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004

C;Accession: S31481

R;Rappold, C.S.J.; Klingmueller, W.

submitted to the EMBL Data Library, January 1993

A;Description: Cloning and sequencing of the recA gene from Enterobacter agglomerans 33

A;Reference number: S31480

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-354 <RAP>

A;Cross-references: UNIPROT:P33037; UNIPARC:UPI000016EC3A; GB:L03291; EMBL:Z19517; NID: C;Genetics:

A;Gene: recA

C;Superfamily: recombination protein recA

C;Keywords: ATP; DNA binding; DNA recombination; DNA repair; nucleotide binding; P-loop

F;67-74/Region: nucleotide-binding motif A (P-loop)

F;141-146/Region: nucleotide-binding motif B

F;73/Binding site: ATP (Lys) #status predicted

Query Match 91.1%; Score 1528; DB 2; Length 354;

[illegible]

```

RESULT 7
S37586
      recombination protein recA - Yersinia pestis
N;Alternate names: recombinase A
C;Species: Yersinia pestis
C;date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C;Accession: S37586
R;Kryukov, V.M.; Suchkov, I.Y.; Sazykin, I.S.; Mishankin, B.N.
  submitted to the EMBL Data Library, October 1993
A;description: Complete nucleotide sequence of Yersinia pestis recA gene.
A;Reference number: S37586
A;Accession: S37586
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-356 <KRY>

```

Query Match	90.9%;	Score 1526;	DB 2;	Length 356;
Best Local Similarity	91.6%;	Pred. No. 3.9e-96;		
Matches 304;	Conservative 15;	Mismatches 13;	Indels 0;	Gaps 0;
Qy	1	AIDENKQKALAAALGQIEKQFGKGSINRLGEDRSDMVKTISTGSLSLDIALGAGGELPMGR	60	
Db	2	AIDENKQKALAAALGQIEKQFGKGSINRLGEDRSDMVETISTGSLSLDIALGAGGELPMGR	61	
Qy	61	IVIIYGPSSGKTTLTLOVIAAAREGKTCAFDAEHALDPIYARKLGVDINDLLCSQPD	120	
Db	62	IVIIYGPSSGKTTLTLOVIAAAREGKTCAFDAEHALDPIYAKKLGVDINDLLCSQPD	121	
Qy	121	TGEQALEICDALARSGAVDIVVDSVAALTPKAEIEGEIGDSHGLAARWMSQAMRKLKAG	180	
Db	122	TGEQALEICDALTSGAVDVIIVDSVAALTPKAEIEGEIGDSHGLAARWMSQAMRKLKAG	181	
Qy	181	NLKQSNLTLLIFINQIRMKIGVMGFNPETTTGGNALKPYASVRLDIRRIGAVKEGENVVGS	240	
Db	182	NLKNNTLLIFINQIRMKIGVMGFNPETTTGGNALKPYASVRLDIRRIGAVKOGDVVVVS	241	
Qy	241	ETRVKVVYKNKIAAPFQKAEQFQILYGEGINFYGELVDLGVKEKLTETKAGAWSYKYGKIGQ	300	

```

Db      242  ETRVYVYVKNKTAAPFKQAEFQILYGEININGELVDLGVKLKLEKAGAWSYYYGDKIGQ 303
Qy      301  GKANATAWLKONPETAKEIEKKKRELLLSNPN 332
Db      302  GKANASYLNKENPANAAELDKLREMLNGCN 333

RESULT 8
RQEBPM
recombination protein recA - Proteus mirabilis
N/Alternate names: recombinase A
C/Species: Proteus mirabilis
C/Date: 30-Sep-1990 #sequence_revision 30-Sep-1990 #text_change 09-Jul-2004
C/Accession: S04606
R/Akaboshi, E.; Yip, M.L.R.; Howard-Flanders, P.
Nucleic Acids Res. 17, 4390, 1989
A/Title: Nucleotide sequence of the recA gene of Proteus mirabilis.
A/Reference number: S04606; MUID:89296502; PMID:2544862
A/Accession: S04606
A/Molecule type: DNA
A/Residues: 1-355 <AKA>
A/Cross-references: UNIPROT:P11406; UNIPARC:UPI000016FPD85; GB:X14870; NID:94561
C/Genetics:
A/Name: recA

```

Query Match	89.4%;	Score 1500;	DB 1;	Length 355;
Best Local Similarity	89.1%;	Pred. No. 2.2e-94;		
Matches 294;	Conservative	22;	Mismatches 14;	Indels 0; Gaps 0

Qy	1	AIDENKQALAAALGQIEKQFGKGSIMRLGEDRSDMVKTISTGSLSLDIALCAGAGLPMGR	60
Db	2	AIDENKQALAAALGQIEKQFGKGSIMRLGEDRSMNVETISTGSLSLDVALCAGAGLPRGR	61
Qy	61	IVEIYGPESSGKTTLTLLQVIAAAQREGKTCAFIDAEHALDPIYARKLGVGDINDLLCSQPD	120
Db	62	IVEIYGPESSGKTTLTLLQVIAASAQREGKICAFIDAEHALDPIYAQKLGVDINDLLCSQPD	121
Qy	121	TGEQALIEICDALARSGADVIVDVSVAALTPKAEIEGIEGSHWGLAARMSQAMRKLKAG	180
Db	122	TGEQALIEICDALARSGADVIVDVSVAALTPKAEIEGIEGSHVGLAARMSQAMRKLKAG	181
Qy	181	NLKQSNLTLLIFINQIRMKIGWMPGNPETTTGNNALKFYASVRLDIRRIGAVKEGNNVWGS	240
Db	182	NLKQSNLTLLIFINQIRMKIGWMPGNPETTTGNNALKFYASVRLDIRRIGSVKNQGEVIGS	241
Qy	241	ETRVRVVKVNIKAAPFKQAEQILLYGEGINFYGELVDLGVKEKLEKAGAWTSYKGEKIGQ	300
Db	242	ETRVRVVKVNVKVAAPFKQAEQIMYGEGINTYGELIDLGVKHKLVEKAGAWTSYNGEKIGQ	301
Qy	301	GKANATAWLKDNPNETAKEIEKKVRELLLSN	330
Db	302	GKANATYLNKHEPPMYNELNTKLFREMLNH	331

RESULT 9
E82310
recA protein VC0543 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C/Species: Vibrio cholerae
C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C/Accession: E82310
R/Heidelberg, J.F.; Eisen, W.C.; Nelson, R.A.; Gwin, M.L.; Dodson, R.; Choudhury, D.; Ermolaeva, V.A.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sell
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A/Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

Qy	62	VEIYP	ESSG	KTTL	QVIA	AAQ	REG	KTC	CA	PI	DA	E	H	A	L	D	P	I	Y	A	R	K	L	G	V	D	I	N	L	L	C	S	O	P	D	T	121																								
Db	61	VEIL	P	ESSG	KTTL	LE	I	A	A	A	Q	R	E	G	K	T	C	A	P	I	D	E	H	A	L	D	P	V	Y	A	K	L	G	V	N	I	D	E	L	L	S	O	P	D	T	120															
Qy	122	GEQ	A	L	E	I	C	D	A	L	A	R	S	G	A	D	V	I	V	D	S	V	A	A	L	T	P	K	A	E	I	G	E	I	G	D	S	H	M	G	L	A	R	M	S	O	A	M	R	K	L	A	G	N	181						
Db	121	GEQ	A	L	E	I	C	D	A	L	A	R	S	G	A	D	V	I	V	D	S	V	A	A	L	T	P	K	A	E	I	G	E	M	G	D	S	H	M	G	L	O	A	R	M	S	O	A	M	R	K	L	T	G	N	180					
Qy	182	L	Q	S	N	T	L	L	I	F	I	N	O	I	R	M	K	I	G	V	M	F	G	N	P	E	T	T	T	G	N	A	L	K	P	Y	A	S	V	R	L	D	I	R	R	I	G	A	V	K	E	G	N	V	V	G	S	E	241		
Db	181	L	Q	S	N	C	M	C	I	F	I	N	O	I	R	M	K	I	G	V	M	F	G	N	P	E	T	T	T	G	N	A	L	K	F	Y	A	S	V	R	L	D	I	R	R	T	G	A	I	K	E	G	E	V	V	G	N	E	240		
Qy	242	T	R	V	K	V	K	N	K	I	A	A	P	P	K	A	S	F	O	I	L	G	E	G	I	N	F	Y	G	E	L	V	D	L	G	V	K	E	L	I	T	E	K	A	G	A	W	S	Y	K	E	K	I	G	O	301					
Db	241	T	R	V	K	V	K	N	K	I	A	A	P	P	K	A	S	T	O	I	M	I	G	F	Y	G	E	L	V	D	L	G	V	K	E	L	I	D	L	G	V	K	H	R	M	E	K	S	G	A	W	S	Y	N	G	D	K	I	G	O	300
Qy	302	K	A	N	A	T	A	L	K	O	N	P	E	T	A	K	E	I	E	K	V	R	E	L	L	S	N	P	N	332																															
Db	301	K	A	N	A	C	K	Y	L	K	E	N	P	E	I	A	K	T	D	K	L	R	E	M	L	N	P	E	N	331																															

Qy 242 TRVVVKNKIAPFPQAQFQIYLEGFINFYGELVDLGVKKEKLI EKAGAWTSYKGEGIKQG 301
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 241 TRIKVVNKKIAPPFKEANTQIMYGQGFNRGEGELIDLGVKHKWVEKSGAWTSYNGDKIKQG 300
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Qy 302 KANATATWLKDNPETAKEIEKKVRELLLSNPN 332
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Db 301 KANACKYLKENPEIAKTILDKKLREMLLNPN 331
 |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 11

JQ1461

recombination protein recA - Vibrio anguillarum

N;Alternate names: recombinase A

C;Species: Vibrio anguillarum

S;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 09-Jul-2004

C;Accession: JQ1461

R;Tolmashky, M.E.; Gammie, A.E.; Crosa, J.H.
Gene 110, 41-48, 1992

A;Title: Characterization of the recA gene of Vibrio anguillarum.

A;Reference number: JQ1461; MUID:92184113; PMID:1544576

A;Accession: JQ1461

A;Molecule type: DNA

A;Residues: 1-348

A;Cross-references: UNIPROT:P26348; UNIPARC:UPI00000133544; GB:M80525; NID:g1551

A;Experimental source: strains 775 and 531A

	Query Match	83.9%; Score 1408.5; DB 2; Length 348;
	Best Local Similarity	83.3%; Pred. No. 3.4e-88;
	Matches 275; Conservative 31; Mismatches 23; Indels 1; Gaps 1	
QY	2 IDENKQKALAAALGQIEKFGKSGIMRLGDSMDVKTTSTGSLDLDIALGAGGLPWGRI	61
	1 MDENKQKALAAALGQIEKFGKSGIMRLGDNRTMDVETSTGSLDLDIALGAGGLPWGRI	60
QY	62 VEIYGPSSGKTTTLTQVIAAARGKTKCAFIDAEHALDPFYARKLGVDINDLLCSQPTT	121
Db	61 VEIYGPSSGKTTTLTLLIAAARGVTKCAFIDAEHALDPFYARKLGVDINDLLVSQPTT	120

[illegible]

||||| i::||| | ::||::||| |||
Db 301 KANACKFLRENPAARAAWALDTKLEMLL-NP 329

RESULT 12
A49929
recombination protein recA - Haemophilus influenzae (strain Rd KW20)
N:Alternate names: rec-1 protein; recombinase A
C:Species: Haemophilus influenzae
C:Date: 11-Nov-1994 #sequence_revision 11-Nov-1994 #text_change 09-Jul-2004
C:Accession: A49929; A64080
R:Zultzy, J.J.; Barcak, G.J.
J. Bacteriol. 175, 7269-7281, 1993
A:Title: Structural organization, nucleotide sequence, and regulation of the Haemophilus influenzae recA gene
A:Reference number: A49929; MUID:94042901; PMID:8228674
A:Accession: A49929
A:Molecule type: DNA
A:Residues: 1-354 <ZUL>
A:Cross-references: UNIPROT:P43705; UNIPARC:UPI00001334E8; GB:L07529; NID:g305381; PIDN:Rd strain KW20
A:Experimental source: Rd strain KW20
R:Flaischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Geoghagen, J.L.; Geoghagen, N.S.M. Science 269, 496-512, 1995
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, A. Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. A:Reference number: A64080; MUID:95350630; PMID:7542800
A:Accession: A64080
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-354 <TIGR>
A:Cross-references: UNIPARC:UPI00001334E8; GB:U32741; GB:L42023; NID:g1573582; PIDN:Rd strain KW20
A:Experimental source: Rd strain KW20
C:Genetics:
A:Function:
A:Description: plays an essential role in homologous recombination, in induction of the C:Superfamily: recombination protein recA
C:Keywords: ATP; DNA binding; DNA recombination; DNA repair; nucleotide binding; P-loop; F:67-74/Region: nucleotide-binding motif A (P-loop)
F:141-146/Region: nucleotide-binding motif B
F:73/Binding site: ATP (lys) #status predicted

Query Match 79.1%; Score 1328; DB 2; Length 354;
Best Local Similarity 76.1%; Pred. No. 1e-82;
Matches 252; Conservative 43; Mismatches 36; Indels 0; Gaps 0;

Qy 1 AIDENKQKALAAALQIEKQFGKSGIMRLGEDRSMVKTITSGSLSLDIALGAGLPMGR 60
Db 2 ATQBEKQKALAAALQIEKQFGKSGIMKLGDTKTLDVESITSGSLGLDVALGIGGLPMGR 61
Qy 61 IVEIYGPSSSGKTTTLTVIAAQAQREGTKCAFIDAEHALDPIYARKLGVDDINLLCSQPD 120
Db 62 IVEIYGPSSSGKTTTLTVIAQAQKAGTKCAFIDAEHALDPIYAAKLGVDVKELFVSPD 121
Qy 121 TGEQALEICDALARGAVDVIIVDSVAALTTPKAEIEGISHGMLAARMWSQAMRKLAG 180
Db 122 NGEQALEICDALVRSGAIDVIIVDSVAALTTPKAEIEGDMGDSHMLQARLMSQALRKLTG 181
Qy 181 NLKQSNLTLLIFINQIRMKIGVFNFGNPETTTGNALKFYASVRLDIRRIGAVKEGENVVGS 240
Db 182 QIKRANCLVFINQIRMKIGVFNFGNPETTTGNALKFYSSVRLDIRRIGAVKEGENVVGS 241
Qy 241 ETRVKVKNKTAAPKQAEFQILYEGEINFGYELVDLGVKEKLIEKAGAWYSYKGEKIQG 300
Db 242 ETRVKVKNKLAAPRPQVDFQILYEGEISKAGELLEGLGVKHLVEKSGAWYSYNGEKIQG 301
Qy 301 KANATATWLKONPETAKIEKKVRELLLSNP 331
Db 302 GKANSRMKWLNNENIEKSDSEARLARLAVANP 332

RESULT 13

JC5198
recombination protein recA - Aeromonas salmonicida
N:Alternate names: recombinase A
C:Species: Aeromonas salmonicida
C:Date: 27-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 09-Jul-2004
C:Accession: JC5198
R:Umelo, E.; Noonan, B.; Trust, T.J.
Gene 175, 133-136, 1996
A:Title: Cloning, characterization and expression of the recA gene of Aeromonas salmonicida
A:Reference number: JC5198; MUID:97074662; PMID:8917089
A:Accession: JC5198
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-353 <UME>
A:Cross-references: UNIPROT:P94190; UNIPARC:UPI00001334C5; GB:U83688; NID:g1785949; PIDN:JC5198
C:Genetics:
A:Gene: recA
C:Function:
A:Description: plays an essential role in homologous recombination, in induction of the C:Superfamily: recombination protein recA
C:Keywords: ATP; DNA binding; DNA recombination; DNA repair; nucleotide binding; P-loop; F:65-72/Region: nucleotide-binding motif A (P-loop)
F:139-144/Region: nucleotide-binding motif B
F:71/Binding site: ATP (lys) #status predicted

Query Match 78.6%; Score 1319; DB 2; Length 353;
Best Local Similarity 79.0%; Pred. No. 4e-82;
Matches 259; Conservative 31; Mismatches 38; Indels 0; Gaps 0;

Qy 2 IDENKQKALAAALQIEKQFGKSGIMRLGEDRSMVKTITSGSLSLDIALGAGLPMGR 61
Db 1 MDQNKQKALAAALQIEKQFGKSGIMLLGDSKTWDIEAISTGSLDVALGIGGLPCGRI 60
Qy 62 VEIYGPSSSGKTTTLTVIAAQAQREGTKCAFIDAEHALDPIYARKLGVDDINLLCSQPD 121
Db 61 VEIYGPSSSGKTTTLTVIAEAAQKGVKCAFIADAEHALDPIYAAKLGVNVVDLLISQSDT 120
Qy 122 GEQALEICDALARGAVDVIIVDSVAALTTPKAEIEGISHGMLAARMWSQAMRKLAGN 181
Db 121 GEQALEICDMLVRNADVVIIVDSVAALTTPKAEIEGEGSHVGLQARLMSQALRKLTAN 180
Qy 182 LKQSNLTLLIFINQIRMKIGVFNFGNPETTTGNALKFYASVRLDIRRIGAVKEGENVVGS 241
Db 181 IKVANCLCIFINQIRMKIGVFNFGSPETTTGNALKFYASVRLDIRRIGAIKEGDEVVGN 240
Qy 242 TRVKVKNKTAAPKQAEFQILYEGEINFGYELVDLGVKEKLIEKAGAWYSYKGEKIQG 301
Db 241 TRVKVKNKVAAPPKQAEFQIFYGVGISKEGELVDLGVKHLIDKAGAWYSYNGEKIQG 300
Qy 302 KANATATWLKONPETAKIEKKVRELLLS 329
Db 301 KANVKKLFTENKVAEVEARLRELLLS 328

RESULT 14
RQPSNA
RecA protein PA3617 [imported] - Pseudomonas aeruginosa (strain PAO1)
N:Alternate names: recombinase A
C:Species: Pseudomonas aeruginosa
C:Date: 30-Sep-1990 #sequence_revision 30-Sep-1990 #text_change 09-Jul-2004
C:Accession: S06265; S10458; A49854; JN0304; G83192
R:Sano, Y.; Kageyama, M.
Mol. Gen. Genet. 208, 412-419, 1987
A:Title: The sequence and function of the recA gene and its protein in Pseudomonas aeruginosa
A:Reference number: S06265; MUID:88038334; PMID:2823059
A:Accession: S06265
A:Molecule type: DNA
A:Residues: 1-346 <SAN1>
A:Cross-references: UNIPROT:P08280; UNIPARC:UPI0000133516; GB:X05691; NID:g45413; PIDN:C
R:Zaitsev, E.N.; Krjukov, V.M.; Kuzmin, N.P.; Alekseev, A.A.; Lanzov, V.A. submitted to the EMBL Data Library, March 1990
A:Description: Restriction polymorphism and nucleotide sequence substitutions in the rec

A;Reference number: S10458
A;Accession: S10458
A;Molecule type: DNA
A;Residues: 1-346 <ZAI>
A;Cross-references: UNIPARC:UPI0000133516; EMBL:X52261; NID:g45383; PIDN:CAA36504.1; PID: J.Sano, Y.
J. Bacteriol. 175, 2451-2454, 1993
A;Title: Role of the recA-related gene adjacent to the recA gene in *Pseudomonas aeruginosa*
A;Reference number: A49854; MUID:93224470; PMID:8468303
A;Accession: A49854
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 313-346 <SAN2>
A;Cross-references: UNIPARC:UPI00001748B1; GB:D13090; NID:g286185
A;Note: sequence extracted from NCBI Backbone (NCBIN:129317, NCBI:P:129318)
R;Kryukov, V.M.; Zaitsev, E.N.; Kouzmin, N.P.; Bayev, A.A.
Bioorg. Khim. 16, 1177-1182, 1990
A;Title: Structure of the recA gene from *Pseudomonas aeruginosa*.
A;Reference number: JN0304; MUID:91182156; PMID:2127886
A;Accession: JN0304
A;Molecule type: DNA
A;Residues: 1-332, 'L', 334-346 <KRY>
A;Cross-references: UNIPARC:UPI00001748B2; GB:X52261; NID:g45383; PIDN:CAA36504.1; PID:G A;Experimental source: strain PM 7
A;Note: the authors translated the codon CTG for residue 333 as Val
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, .; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic patho
A;Reference number: A82950; MUID:2043737; PMID:10984043
A;Accession: G83192
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-346 <STO>
A;Cross-references: UNIPARC:UPI0000133516; GB:AE004782; GB:AE004091; NID:g9949772; PIDN: A;Experimental source: strain PA01
C;Genetics:
A;Gene: recA; PA3617
C;Function:
A;Description: plays an essential role in homologous recombination, in induction of the
C;Superfamily: recombination protein recA
C;Keywords: ATP; DNA binding; DNA recombination; DNA repair; nucleotide binding; P-loop;
F;65-72/Region: nucleotide-binding motif A (P-loop)
F;139-144/Region: nucleotide-binding motif B
F;71/Binding site: ATP (Lys) #status predicted
Query Match 75.6%; Score 1268; DB 1; Length 346;
Best Local Similarity 74.4%; Pred. No. 1.1e-78;
Matches 244; Conservative 42; Mismatches 42; Indels 0; Gaps 0;
Qy 2 IDENKOKALAAALGOIEKQFGKSGIMRLGEDRSMDVKTISTGSLDIALGAGGLPMGRI 61
Db 1 MDENKKALAAALGOIEKQFGKSGIMRLGEDRSMDVKTISTGSLDIALGAGGLPMGRI 60
Qy 62 VEIYGPESSGKTTTLTQVIAAAQREGKTCAPIDAEHALDPIYARKLGVYDIDNLLCSQPD 121
Db 61 VEIYGPESSGKTTTLTQVIAAAQREGKTCAPIDAEHALDPIYARKLGVYDIDNLLCSQPD 120
Qy 122 GEOALETCDALARGAVDVIIVDSVAALTPKAEITEGISHMGLAARMMSQAMRKLGN 181
Db 121 GEOALETCDMLVRNVDVIIVDSVAALTPKAEITEGISHMGLAARMMSQAMRKLGN 180
Qy 182 LKQSNLTLLIFNQIRMKIGVFNFGNPTTTCGNALKFYASVRLDIRRTGAVKEGENVVGS 241
Db 181 IQNANCLVIFNQIRMKIGVFNFGNPTTTCGNALKFYASVRLDIRRTGAVKEGENVVGS 240
Qy 242 TRVKVKNKIAAPFKQAEFQILYGEINFGYELVDLGVKEKLEKAGATSYKGEKIQG 301
Db 241 TRVKVKNKVPFPRQAEFQILYKGIRNGEIIDLGVQLGLLEKSGAWSYQSGKIQG 300
Qy 302 KANATAWLKONPETAKEIEKKVRELLS 329
Db 301 KANATAWLKONPETAKEIEKKVRELLS 329

Search completed: February 16, 2006, 01:24:57
Job time : 26.5 secs

Db 301 KANAAKYLEDNPEIGSVLEKTRDQLLA 328
RESULT 15
JN0321
recombination protein recA - *Azotobacter vinelandii*
N;Alternate names: recombinase A
C;Species: *Azotobacter vinelandii*
C;Date: 09-Oct-1992 #sequence_revision 09-Oct-1992 #text_change 09-Jul-2004
C;Accession: JN0321
R;Kenkatesh, T.V.; Das, H.K.
Gene 113, 47-53, 1992
A;Title: The *Azotobacter vinelandii* recA gene: sequence analysis and regulation of exp
A;Reference number: JN0321; MUID:9225347; PMID:1563632
A;Accession: JN0321
A;Molecule type: DNA
A;Residues: 1-349 <KEN>
A;Cross-references: UNIPROT:P29246; UNIPARC:UPI00001334CC
A;Note: the authors translated the codon GAC for residue 129 as Val and AAC for residue
C;Genetics:
A;Gene: recA
C;Function:
A;Description: plays an essential role in homologous recombination, in induction of the
C;Superfamily: recombination protein recA
C;Keywords: ATP; DNA binding; DNA recombination; DNA repair; DNA replication; nucleotid
F;65-72/Region: nucleotide-binding motif A (P-loop)
F;139-144/Region: nucleotide-binding motif B
F;71/Binding site: ATP (Lys) #status predicted

Query Match 75.4%; Score 1265; DB 2; Length 349;
Best Local Similarity 73.5%; Pred. No. 1.8e-78;
Matches 244; Conservative 41; Mismatches 47; Indels 0; Gaps 0;
Qy 2 IDENKOKALAAALGOIEKQFGKSGIMRLGEDRSMDVKTISTGSLDIALGAGGLPMGRI 61
Db 1 MDENKKALAAALGOIEKQFGKSGIMRLGEDRSMDVKTISTGSLDIALGAGGLPMGRI 60
Qy 62 VEIYGPESSGKTTTLTQVIAAAQREGKTCAPIDAEHALDPIYARKLGVYDIDNLLCSQPD 121
Db 61 VEIYGPESSGKTTTLTQVIAAAQREGKTCAPIDAEHALDPIYARKLGVYDIDNLLCSQPD 120
Qy 122 GEOALETCDALARGAVDVIIVDSVAALTPKAEITEGISHMGLAARMMSQAMRKLGN 181
Db 121 GEOALETCDMLVRNVDVIIVDSVAALTPKAEITEGISHMGLAARMMSQAMRKLGN 180
Qy 182 LKQSNLTLLIFNQIRMKIGVFNFGNPTTTCGNALKFYASVRLDIRRTGAVKEGENVVGS 241
Db 181 IQNANCLVIFNQIRMKIGVFNFGNPTTTCGNALKFYASVRLDIRRTGAVKEGENVVGS 240
Qy 242 TRVKVKNKIAAPFKQAEFQILYGEINFGYELVDLGVKEKLEKAGATSYKGEKIQG 301
Db 241 TRVKVKNKVPFPRQAEFQILYKGIRNGEIIDLGVQLGLLEKSGAWSYQSGKIQG 300
Qy 302 KANATAWLKONPETAKEIEKKVRELLSNPNS 333
Db 301 KANAAKYLEDNPEVAAAWEKSIQDLAAPAS 332

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.
OM protein - protein search, using sw model
Run on: February 16, 2006, 01:14:35 ; Search time 152 Seconds
(without alignments)
1554.947 Million cell updates/sec
Title: US-10-733-782-3
Perfect score: 1678
Sequence: 1 AIDENKOKALAAALGQIEKQ.....AKEIEKKVRELLSNPNSTP 335

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 2166443 seqs, 705528306 residues
Total number of hits satisfying chosen parameters: 2166443
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
Database : UniProt_05.80.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1674	99.8	352	1 RECA_ECO57	P0a798 escherichia
2	1674	99.8	352	1 RECA_ECOL6	P0a797 escherichia
3	1674	99.8	352	1 RECA_ECOLI	P0a796 escherichia
4	1674	99.8	352	1 RECA_SHIFL	P0a799 shigella fl
5	1667	99.3	352	1 RECA_SHISO	Q9zff6 shigella so
6	1637	97.6	352	1 RECA_SALTY	P65978 salmonella
7	1637	97.6	352	1 RECA_SALTY	P65977 salmonella
8	1637	97.4	353	2 Q57KU4 SALCH	Q57ku4 salmonella
9	1635	97.4	353	2 Q5PFI5 SALPA	Q5pfi5 salmonella
10	1543	92.0	355	1 RECA_PHOLL	Q7n7a6 photorhabdu
11	1533	91.4	357	1 RECA_ERWCT	Q6dl88 erwinia car
12	1530	91.2	355	1 RECA_YERPE	P37858 yersinia pe
13	1530	91.2	356	1 RECA_YERPS	Q6e070 yersinia ps
14	1528	91.1	353	1 RECA_ENTAG	P33037 enterobacte
15	1523	90.8	353	1 RECA_SERMA	P17479 serratia ma
16	1520	90.6	358	1 RECA_XENBV	P96185 xenorhabdu
17	1519	90.5	355	1 RECA_SODGL	P62220 sodalis glo
18	1509	89.9	355	2 Q9YI31_9ENTR	Q9y131 primary end
19	1500	89.4	354	1 RECA_PROMI	P11406 proteus mir
20	1497	89.2	342	1 RECA_ERWCA	P26344 erwinia car
21	1497	89.2	357	1 RECA_SHEON	Q8eb80 shewanella
22	1496	89.2	358	1 RECA_XENNE	Q9x5p5 xenorhabdu
23	1482	88.3	325	1 RECA_PROVU	P26346 proteus vul
24	1433	85.4	354	1 RECA_VIBCH	P45383 vibrio chol
25	1425	84.9	347	1 RECA_VIBPA	Q871r1 vibrio para
26	1414	84.3	348	1 RECA_VIBNA	Q6xz07 vibrio natr
27	1414	84.3	349	1 RECA_VIBVU	Q8dc51 vibrio vuln
28	1414	84.3	349	1 RECA_VIBVY	Q7mhr4 vibrio vuln
29	1412	84.1	348	2 Q5E7G6_VIBFI	Q5e7g6 vibrio fisc
30	1408.5	83.9	348	1 RECA_VIBAN	P26348 vibrio angu
31	1400	83.4	352	1 RECA_PHOPR	Q61mu2 photobacter

32	1396	83.2	346	2 Q9S4R6_VIBCH	Q9s4r6 vibrio chol
33	1396	83.2	346	2 Q9R2W9_VIBCH	Q9r2w9 vibrio chol
34	1349	80.4	354	1 RECA_PASMU	P95526 pasteurella
35	1328	79.1	354	1 RECA_HAEIN	P43705 haemophilus
36	1328	79.1	354	2 Q4QM72_HAE18	Q4qmv2 haemophilus
37	1323	78.8	348	1 RECA_PSEOL	Q9t9u2 pseudomonas
38	1323	78.8	372	2 Q5QB8_IDILO	Q5qub8 idiomarina
39	1319	78.6	353	1 RECA_AERSA	P94190 aeromonas s
40	1315	78.4	351	1 RECA_MANSB	Q65qbo manheimia
41	1313	78.2	331	1 RECA_WIGBR	Q8d2w7 wiggleswort
42	1312	78.2	352	1 RECA_ACTAC	Q9itp9 actinobacil
43	1284	76.5	368	1 RECA_PASHA	Q9rny0 pasteurella
44	1276	76.0	347	2 Q6EV36_FSEST	Q6ev36 pseudomonas
45	1275	76.0	349	2 Q4IXX9_AZOVI	Q4ixx9 azotobacter

ALIGNMENTS

RESULT 1

ID RECA_ECO57 STANDARD; PRT; 352 AA.
AC P0A7G8: P03017; P26347; P78213;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE RECA protein (Recombinase A).
GN Name=recA; OrderedLocName=z4002, ECa3556;
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;
RX MEDLINE=12074935; PubMed=11206551; DOI=10.1038/35054089;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
Welch R.A., Blattner F.R.,
RL Nature 409:529-533(2001).
RL [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=O157:H7 / Sakai / RIMD 0509952 / EHEC;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
Kuhara S., Shiba T., Hattori M., Shinagawa H.;
*Complete genome sequence of enterohemorrhagic Escherichia coli
O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
CC -!- FUNCTION: Can catalyze the hydrolysis of ATP in the presence of
single-stranded DNA, the ATP-dependent uptake of single-stranded
DNA by duplex DNA, and the ATP-dependent hybridization of
homologous single-stranded DNAs. It interacts with lexA causing
its activation and leading to its autocatalytic cleavage (by
similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (by similarity).
CC -!- SIMILARITY: Belongs to the recA family.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
CC EMBL; AB005174; AAC57804.1; -; Genomic_DNA.
CC EMBL; BA000007; BAB36979.1; -; Genomic_DNA.
CC PIR; D91073; D91073.

```
DR HAMAP; MF_00268; -; 1.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001553; RecA.
DR Pfam; PF00154; RecA; 1.
DR PRINTS; PD00142; RecA.
DR ProDom; PD00229; RecA; 1.
DR SMART; SM00382; AAA; 1.
DR TIGRfams; TIGR02012; tigrfam_reca; 1.
DR PROSITE; PS00321; RECA_1; 1.
DR PROSITE; PS0162; RECA_2; 1.
DR PROSITE; PS0163; RECA_3; 1.
KW ATP-binding; Complete proteome; DNA damage; DNA recombination;
KW DNA repair; DNA-binding; Nucleotide-binding; SOS response.
FT INIT_MET 0
FT NP_BIND 66 73
FT CONFLICT 139 139 D -> N (in Ref. 1).
SQ SEQUENCE 352 AA; 37842 MW; 989B02378EC4A402 CRC64;

Query Match          99.8%; Score 1674; DB 1; Length 352;
Best Local Similarity 99.7%; Pred. No. 1.6e-103;
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVKTISTGSLSLDIALGAGGLPMGR 60
DB 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVKTISTGSLSLDIALGAGGLPMGR 60
QY 61 IVEIYGPSSGKTTLTQVIAAAQREGKTCFADAEHALDPIYARKLGVLDNLLCSQPD 120
DB 61 IVEIYGPSSGKTTLTQVIAAAQREGKTCFADAEHALDPIYARKLGVLDNLLCSQPD 120
QY 121 TGEQALEICDALARGSDVIVVDSVAALTPKAEIEGEIGDSHMGLAARMWSQAMRKLKAG 180
DB 121 TGEQALEICDALARGSDVIVVDSVAALTPKAEIEGEIGDSHMGLAARMWSQAMRKLKAG 180
QY 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
DB 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
QY 241 ETRVVKVVKNTIAAPFKQAEFQILYEGEINFGYELVDLGKVEKLEKAGWYSYKGEKIQG 300
DB 241 ETRVVKVVKNTIAAPFKQAEFQILYEGEINFGYELVDLGKVEKLEKAGWYSYKGEKIQG 300
QY 301 GKANATAWLKNPETAKEIEKKVRELLLSNPSTP 335
DB 301 GKANATAWLKNPETAKEIEKKVRELLLSNPSTP 335
```

RESULT 2

```
RECA_EC0L6          STANDARD;          PRT;          352 AA.
AC POA7G6; P03017; P26347; P78213;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE RecA protein (Recombinase A).
GN Name=recA; OrderedLocNames=c3253;
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=O6:H1 / CFT073 / ATCC 700928 / UPEC;
RX MEDLINE=2238234; PubMed=12471157; DOI=10.1073/pnas.252529799;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Raeko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
CC -!- FUNCTION: Can catalyze the hydrolysis of ATP in the presence of
CC single-stranded DNA, the ATP-dependent uptake of single-stranded
```

```
CC DNA by duplex DNA, and the ATP-dependent hybridization of
CC homologous single-stranded DNAs. It interacts with lexA causing
CC its activation and leading to its autocatalytic cleavage (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the recA family.
CC
CC This Swiss-Prot entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; AE016765; AAN81704.1; -; Genomic_DNA.
DR HAMAP; MF_00268; -; 1.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001553; RecA.
DR Pfam; PF00154; RecA; 1.
DR PRINTS; PD00142; RecA.
DR ProDom; PD00229; RecA; 1.
DR SMART; SM00382; AAA; 1.
DR TIGRfams; TIGR02012; tigrfam_reca; 1.
DR PROSITE; PS00321; RECA_1; 1.
DR PROSITE; PS0162; RECA_2; 1.
DR PROSITE; PS0163; RECA_3; 1.
KW ATP-binding; Complete proteome; DNA damage; DNA recombination;
KW DNA repair; DNA-binding; Nucleotide-binding; SOS response.
FT INIT_MET 0
FT NP_BIND 66 73
FT CONFLICT 139 139 D -> N (in Ref. 1).
SQ SEQUENCE 352 AA; 37842 MW; 989B02378EC4A402 CRC64;

Query Match          99.8%; Score 1674; DB 1; Length 352;
Best Local Similarity 99.7%; Pred. No. 1.6e-103;
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVKTISTGSLSLDIALGAGGLPMGR 60
DB 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVKTISTGSLSLDIALGAGGLPMGR 60
QY 61 IVEIYGPSSGKTTLTQVIAAAQREGKTCFADAEHALDPIYARKLGVLDNLLCSQPD 120
DB 61 IVEIYGPSSGKTTLTQVIAAAQREGKTCFADAEHALDPIYARKLGVLDNLLCSQPD 120
QY 121 TGEQALEICDALARGSDVIVVDSVAALTPKAEIEGEIGDSHMGLAARMWSQAMRKLKAG 180
DB 121 TGEQALEICDALARGSDVIVVDSVAALTPKAEIEGEIGDSHMGLAARMWSQAMRKLKAG 180
QY 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
DB 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
QY 241 ETRVVKVVKNTIAAPFKQAEFQILYEGEINFGYELVDLGKVEKLEKAGWYSYKGEKIQG 300
DB 241 ETRVVKVVKNTIAAPFKQAEFQILYEGEINFGYELVDLGKVEKLEKAGWYSYKGEKIQG 300
QY 301 GKANATAWLKNPETAKEIEKKVRELLLSNPSTP 335
DB 301 GKANATAWLKNPETAKEIEKKVRELLLSNPSTP 335

RESULT 3
RECA_EC0L1          STANDARD;          PRT;          352 AA.
AC POA7G6; P03017; P26347; P78213;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE RecA protein (Recombinase A).
GN Name=recA; Synonyms=lexB, rech, rnmB, tif, umuB, zab;
GN OrderedLocNames=b2699;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
```

OX NCBI_TaxID=562;
RN [1]
RP NUCLEOTIDE SEQUENCE, AND PROTEIN SEQUENCE OF 1-5.
RX MEDLINE=80145618; PubMed=6244554;
RA Horii T., Ogawa T., Ogawa H.;
RT "Organization of the recA gene of *Escherichia coli*.";
RN Proc. Natl. Acad. Sci. U.S.A. 77:313-317(1980).
[2]
RP NUCLEOTIDE SEQUENCE, AND PROTEIN SEQUENCE OF 1-2.
RX MEDLINE=80234673; PubMed=6930655;
RA Sancar A., Stachelle C., Konigsberg W., Rupp W.D.;
RT "Sequences of the recA gene and protein.";
RN Proc. Natl. Acad. Sci. U.S.A. 77:2611-2615(1980).
[3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=91109725; PubMed=2274037;
RA Zhao X.J., McEntee K.;
RT "DNA sequence analysis of the recA genes from *Proteus vulgaris*,
RN *Erwinia carotovora*, *Shigella flexneri* and *Escherichia coli* B/r.";
RN Mol. Gen. Genet. 222:369-376(1990).
[4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX STRAIN=K12 / MG1655;
RN MEDLINE=97426617; PubMed=9278503; DOI=10.1126/science.277.5331.1453;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RN Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of *Escherichia coli* K-12.";
RN Science 277:1453-1474(1997).
[5]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=K12;
RN MEDLINE=97349980; PubMed=9205837;
RA Yamamoto Y., Aiba H., Baba T., Hayaehi K., Inada T., Isono K.,
RN Ichio T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsuhashi N.,
RA Mizobuchi K., Mori H., Nakade S., Nakamura Y., Nishimoto H.,
RN Ohnishi T., Oyama S., Saito N., Sampel G., Satoh Y., Sivasubram S.,
RA Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,
RN Yamagata S., Horiuchi T.;
RT "Construction of a contiguous 874-kb sequence of the *Escherichia coli*-
RN K12 genome corresponding to 50.0-68.8 min on the linkage map and
RT analysis of its sequence features.";
RN DNA Res. 4:91-113(1997).
[6]
RP PROTEIN SEQUENCE OF 89-106 AND 178-183.
RX MEDLINE=95255284; PubMed=7737176;
RA Morimatsu K., Horii T.;
RN "The DNA-binding site of the RecA protein. Photochemical cross-linking
RT of Tyr103 to single-stranded DNA.";
RN Eur. J. Biochem. 228:772-778(1995).
[7]
RP PROTEIN SEQUENCE OF 186-193.
RX MEDLINE=96067680; PubMed=7588783;
RA Gardner R.V., Voloshin O.N., Camerini-Otero R.D.;
RN "The identification of the single-stranded DNA-binding domain of the
RT *Escherichia coli* RecA protein.";
RN Eur. J. Biochem. 233:419-425(1995).
[8]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
RX MEDLINE=92114994; PubMed=1731246; DOI=10.1038/355318a0;
RA Story R.M., Weber I.T., Steitz T.A.;
RN "The structure of the *E. coli* recA protein monomer and polymer.";
RN Nature 355:318-325(1992).
[9]
RP ERRATUM.
RA Story R.M., Weber I.T., Steitz T.A.;
RN Nature 355:567-567(1992).
[10]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
RX MEDLINE=97185905; PubMed=9033586;
RA Yu X., Egelman E.H.;
RT "The RecA hexamer is a structural homologue of ring helicases.";

RL Nat. Struct. Biol. 4:101-104(1997).
RN [11]
RP STRUCTURE OF ATP-BINDING FOLD.
RX MEDLINE=92115005; PubMed=1731253; DOI=10.1038/355374a0;
RA Story R.M., Steitz T.A.;
RT "Structure of the recA protein-ADP complex.";
RN Nature 355:374-376(1992).
CC -!- FUNCTION: Can catalyze the hydrolysis of ATP in the presence of
CC single-stranded DNA, the ATP-dependent uptake of single-stranded
CC DNA by duplex DNA, and the ATP-dependent hybridization of
CC homologous single-stranded DNAs. It interacts with LexA causing
CC its activation and leading to its autocatalytic cleavage.
CC -!- INTERACTION:
CC P0A6G7:clpP; NbExp=1; IntAct=EBI-370331; EBI-370625;
CC P08622:dnaJ; NbExp=1; IntAct=EBI-370331; EBI-545285;
CC P77488:dxs; NbExp=1; IntAct=EBI-370331; EBI-549927;
CC P06138:ftsZ; NbExp=1; IntAct=EBI-370331; EBI-370963;
CC P11120:glmM; NbExp=1; IntAct=EBI-370331; EBI-370683;
CC P09097:gyrA; NbExp=1; IntAct=EBI-370331; EBI-547129;
CC P0A6H5:holU; NbExp=1; IntAct=EBI-370331; EBI-369317;
CC P0A817:metK; NbExp=1; IntAct=EBI-370331; EBI-546295;
CC P13519:metB; NbExp=1; IntAct=EBI-370331; EBI-371008;
CC P09152:narG; NbExp=1; IntAct=EBI-370331; EBI-547248;
CC P05055:ppp; NbExp=1; IntAct=EBI-370331; EBI-548080;
CC P0A7B9:pyrH; NbExp=1; IntAct=EBI-370331; EBI-370182;
CC P03842:rph; NbExp=1; IntAct=EBI-370331; EBI-557453;
CC P36979:vfGB; NbExp=1; IntAct=EBI-370331; EBI-559071;
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- INDUCTION: In response to low temperature. Sensitive to
CC temperature through changes in the linking number of the DNA.
CC -!- SIMILARITY: Belongs to the recA family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; V00328; CAA23618.1; -; Genomic DNA.
CC EMBL; U00096; AAC75741.1; -; Genomic DNA.
CC EMBL; D90892; BAA16561.1; -; Genomic DNA.
CC F01; G65049; R0ECA.
CC P01; 1AA3; NMR; @=-.
CC PDB; 1N03; EM; A/B/C/D/E/F/G=1-352.
CC PDB; 1REA; X-ray; @=-.
CC PDB; 1U94; X-ray; A=1-352.
CC PDB; 1U98; X-ray; A=1-352.
CC PDB; 1U99; X-ray; A=1-352.
CC PDB; 1XMS; X-ray; A=1-352.
CC PDB; 1XWV; X-ray; A=1-352.
CC PDB; 2REB; X-ray; @=-.
CC PDB; 2REC; EM; A/B/C/D/E/F=1-352.
CC IntAct; P0A7G6; -.
CC SWISS-2DPAGE; P0A7G6; COLI.
CC ECO2DBASE; C039.3; 6TH EDITION.
CC EcoBASE; EB0816; -.
CC EcoGene; EG10823; recA.
CC HAMAP; MF_00268; -; 1.
CC InterPro; IPR003593; AAA_ATPase.
CC InterPro; IPR001553; RecA.
CC Pfam; PF00154; RecA; 1.
CC PRINTS; PR00142; RECA.
CC ProDom; PD000229; RecA; 1.
CC SMART; SM00382; AAA; 1.
CC TIGRfams; TIGR02012; tigrfam_recA; 1.
CC PROSITE; PS00321; RECA_1; 1.
CC PROSITE; PS0162; RECA_2; 1.
CC PROSITE; PS0163; RECA_3; 1.
CC 3D-structure; ATP-binding; Complete proteome;
CC Direct protein sequencing; DNA damage; DNA recombination; DNA repair;
CC DNA-binding; Nucleotide-binding; SOS response.
FT INIT MET 0 0
FT NP_BIND 66 73 ATP.

FT	CONFLICT	112	112	D -> E (in Ref. 5).	
FT	CONFLICT	190	190	Missing (in Ref. 7).	
FT	TURN	4	4		
FT	HELIX	5	21		
FT	HELIX	23	25		
FT	TURN	29	30		
FT	STRAND	39	40		
FT	HELIX	45	50		
FT	TURN	51	51		
FT	STRAND	56	57		
FT	TURN	58	59		
FT	STRAND	61	65		
FT	TURN	68	69		
FT	HELIX	72	85		
FT	TURN	86	87		
FT	STRAND	90	94		
FT	HELIX	101	106		
FT	TURN	107	108		
FT	HELIX	111	113		
FT	STRAND	115	117		
FT	STRAND	122	135		
FT	STRAND	140	144		
FT	HELIX	146	148		
FT	HELIX	152	155		
FT	HELIX	166	185		
FT	TURN	186	186		
FT	STRAND	188	193		
FT	HELIX	213	218		
FT	STRAND	221	233		
FT	TURN	234	235		
FT	STRAND	236	249		
FT	STRAND	257	263		
FT	TURN	264	266		
FT	STRAND	267	268		
FT	HELIX	270	280		
FT	TURN	281	282		
FT	STRAND	285	287		
FT	TURN	288	289		
FT	STRAND	290	293		
FT	TURN	294	295		
FT	STRAND	296	300		
FT	HELIX	301	311		
FT	HELIX	313	327		
Query Match 99.8%; Score 1674; DB 1; Length 352;					
Best Local Similarity 99.7%; Pred. No. 1.6e-103;					
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;					
Qy	1	AIDENKQKALAAALGOIEKQFGKGSIMRLGEDRSMDVKTISTGSLSLDIALGAGGLPMGR	60		
Db	1	AIDENKQKALAAALGOIEKQFGKGSIMRLGEDRSMDVETISTGSLSLDIALGAGGLPMGR	60		
Qy	61	IVEIYGPSSGKTTLTQVIAAAREGKTCAPIDAEHALDPIYARKLGVDDNLLCSQPD	120		
Db	61	IVEIYGPSSGKTTLTQVIAAAREGKTCAPIDAEHALDPIYARKLGVDDNLLCSQPD	120		
Qy	121	TGEQALEICDALSARGVDVIVDSVAALTPKAEIEGIBGSHMGLAARMMSQAMKLAG	180		
Db	121	TGEQALEICDALSARGVDVIVDSVAALTPKAEIEGIBGSHMGLAARMMSQAMKLAG	180		
Qy	181	NLKQSNLTLLIFINQIRMKIGWFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS	240		
Db	181	NLKQSNLTLLIFINQIRMKIGWFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS	240		
Qy	241	ETRVKVNKKAAPKQAEFQILYCEGINFYGELVDLGVKEKLEIKAGAWTSYKGEKIQ	300		
Db	241	ETRVKVNKKAAPKQAEFQILYCEGINFYGELVDLGVKEKLEIKAGAWTSYKGEKIQ	300		
Qy	301	GKANATAWLKDNPETAKEIEKKVRELLLSNPNSTP	335		
Db	301	GKANATAWLKDNPETAKEIEKKVRELLLSNPNSTP	335		

RESULT 4
RECA_SHIFL STANDARD; PRT; 352 AA.
AC POA7G9; P03017; P26347; P78213;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE RecA protein (Recombinase A)
GN Name=recA; OrderedLocustNames=SF2722, S2913;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=BS12 / Serotype 2a;
RX MEDLINE=91109725; PubMed=2274037;
RA Zhao X.J., McEntee K.;
RT "DNA sequence analysis of the recA genes from Proteus vulgaris,
RL Erwinia carotovora, Shigella flexneri and Escherichia coli B/r.";
RN Mol. Gen. Genet. 222:369-376(1990).
[2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=301 / Serotype 2a;
RX MEDLINE=2272406; PubMed=12384590; DOI=10.1093/nar/gkf566;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
through comparison with genomes of Escherichia coli K12 and O157.";
RL Nucleic Acids Res. 30:4432-4441(2002).
[3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=2457T / ATCC 700930 / Serotype 2a;
RX MEDLINE=22590274; PubMed=12704152;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of Shigella
flexneri serotype 2a strain 2457T.";
RL Infect. Immun. 71:2775-2786(2003).
CC -!- FUNCTION: Can catalyze the hydrolysis of ATP in the presence of
CC single-stranded DNA, the ATP-dependent uptake of single-stranded
CC DNA by duplex DNA, and the ATP-dependent hybridization of
CC homologous single-stranded DNAs. It interacts with LexA causing
CC its activation and leading to its autocatalytic cleavage (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the recA family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; X55553; CAB56806.1; -; Genomic DNA.
CC EMBL; AE005674; AA044214.1; ALT_INIT; Genomic DNA.
CC EMBL; AE016987; AAP18040.1; -; Genomic DNA.
CC HAMAP; MF_00268; -; 1.
CC InterPro; IPR003593; AAA_ATPase.
CC InterPro; IPR001553; RecA.
CC Pfam; PF00154; RecA; 1.
CC PRINTS; PR00142; RecA; 1.
CC PRODOM; PD000229; RecA; 1.
CC SMART; SM00382; AAA; 1.
CC TIGRFAMs; TIGR02012; tigrfam_recA; 1.
CC PROSITE; PS00321; RECA_1; 1.
CC PROSITE; PS50162; RECA_2; 1.

```

DR PROSITE; PS50163; RECA_3; 1.
KW ATP-binding; Complete proteome; DNA damage; DNA recombination;
KW DNA repair; DNA-binding; Nucleotide-binding; SOS response.
DR INIT MET 0 By similarity.
FT NP BIND 66 73 ATP (By similarity).
SQ SEQUENCE 352 AA; 37842 MW; 989B02378ECA4A02 CRC64;

Query Match 99.8%; Score 1674; DB 1; Length 352;
Best Local Similarity 99.7%; Pred. No. 1.6e-103;
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AIDENKOKALAAALGOIEKQFGKGSIMRLGSDRSMDVKTISTGSLSLDIALGAGLPMGR 60
DB 1 AIDENKOKALAAALGOIEKQFGKGSIMRLGSDRSMDVKTISTGSLSLDIALGAGLPMGR 60
QY 61 IVEIYGPESSGKTTLTQVIAAAQREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
DB 61 IVEIYGPESSGKTTLTQVIAAAQREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
QY 121 TGEQALEICDALARSGADVIVVDSVAALTPKAEIEGIGDSHMGLAARMMSQAMRKLKAG 180
DB 121 TGEQALEICDALARSGADVIVVDSVAALTPKAEIEGIGDSHMGLAARMMSQAMRKLKAG 180
QY 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
DB 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
QY 241 ETRVKVVKNTAAAPFKQAEFQILYGEINFGYELVDLGVEKLEKAGAWTSYKGEKIGQ 300
DB 241 ETRVKVVKNTAAAPFKQAEFQILYGEINFGYELVDLGVEKLEKAGAWTSYKGEKIGQ 300
QY 301 GKANATAWLKDNPETAKEIEKKVRELLLSNPNSPT 335
DB 301 GKANATAWLKDNPETAKEIEKKVRELLLSNPNSPT 335

RESULT 5
RECA_SHISO STANDARD; PRT; 352 AA.
AC Q9ZFF6;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE RecA protein (Recombinase A).
GN Name=recA;
OS Shigella sonnei.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=624;
[1]
RN NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RC STRAIN=KNTH104S;
RA Park Y.C., Shin H.J., Kim Y.C.;
RT "Cloning and nucleotide sequence of the recA gene from Shigella sonnei
KNTH104S."
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Can catalyze the hydrolysis of ATP in the presence of
CC single-stranded DNA, the ATP-dependent uptake of single-stranded
CC DNA by duplex DNA, and the ATP-dependent hybridization of
CC homologous single-stranded DNAs. It interacts with lexA causing
CC its activation and leading to its autocatalytic cleavage.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the recA family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AF101227; AAC72856.1; -; Genomic_DNA.
CC HSP; P03017; 2REB.
CC SMR; Q9ZFF6; 3-328.

```

```

DR HAVAP; MF_00268; -; 1.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001553; RecA.
DR Pfam; PF00154; RecA; 1.
DR PRINTS; PR00142; RECA.
DR ProDom; PD000229; RecA; 1.
DR SMART; SM00382; AAA; 1.
DR TIGRFAMs; TIGR02012; tigrfam_recA; 1.
DR PROSITE; PS00321; RECA_1; 1.
DR PROSITE; PS01621; RECA_2; 1.
DR PROSITE; PS0163; RECA_3; 1.
KW ATP-binding; DNA damage; DNA recombination; DNA repair; DNA-binding;
KW Nucleotide-binding; SOS response.
DR INIT MET 0 By similarity.
FT NP BIND 66 73 ATP (By similarity).
SQ SEQUENCE 352 AA; 37799 MW; BEC50231893AA40C CRC64;

Query Match 99.3%; Score 1667; DB 1; Length 352;
Best Local Similarity 99.4%; Pred. No. 4.6e-103;
Matches 333; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 AIDENKOKALAAALGOIEKQFGKGSIMRLGSDRSMDVKTISTGSLSLDIALGAGLPMGR 60
DB 1 AIDENKOKALAAALGOIEKQFGKGSIMRLGSDRSMDVKTISTGSLSLDIALGAGLPMGR 60
QY 61 IVEIYGPESSGKTTLTQVIAAAQREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
DB 61 IVEIYGPESSGKTTLTQVIAAAQREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
QY 121 TGEQALEICDALARSGADVIVVDSVAALTPKAEIEGIGDSHMGLAARMMSQAMRKLKAG 180
DB 121 TGEQALEICDALARSGADVIVVDSVAALTPKAEIEGIGDSHMGLAARMMSQAMRKLKAG 180
QY 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
DB 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
QY 241 ETRVKVVKNTAAAPFKQAEFQILYGEINFGYELVDLGVEKLEKAGAWTSYKGEKIGQ 300
DB 241 ETRVKVVKNTAAAPFKQAEFQILYGEINFGYELVDLGVEKLEKAGAWTSYKGEKIGQ 300
QY 301 GKANATAWLKDNPETAKEIEKKVRELLLSNPNSPT 335
DB 301 GKANATAWLKDNPETAKEIEKKVRELLLSNPNSPT 335

RESULT 6
RECA_SALTI STANDARD; PRT; 352 AA.
AC P65978; Q8XET0;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE RecA protein (Recombinase A).
GN Name=recA; OrderedLocusNames=STY2950, t2730;
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=601;
[1]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=CT18;
RC MEDLINE=21534947; PubMed=11677608; DOI=10.1038/35101607;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C.M., Mungall K.L., Bentley S.D., Holden M.T.G., Sebatina M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dow L., White N., Farrar J.,
RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holtroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M.A., Rutherford K.M., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrrell B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
enterica serovar Typhi CT18."

```



```
RL Nature 413:848-852 (2001).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Ty2 / ATCC 700931;
RX MEDLINE=22531367; PubMed=12644504;
RY DOI=10.1128/JB.185.7.2330-2337.2003;
RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RB Burland V., Kodoyanni V., Schwartz D.C., Blattner F.R.,
RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
and Cr18.";
RL J. Bacteriol. 185:2330-2337 (2003).
CC -!- FUNCTION: Can catalyze the hydrolysis of ATP in the presence of
CC single-stranded DNA, the ATP-dependent uptake of single-stranded
CC DNA by duplex DNA, and the ATP-dependent hybridization of
CC homologous single-stranded DNAs. It interacts with LexA causing
CC its activation and leading to its autocatalytic cleavage (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the recA family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; AL627276; CAD05935.1; -; Genomic DNA.
DR EMBL; AB016843; AA070291.1; -; Genomic_DNA.
DR HSSP; P03017; 2REB.
DR SMR; P65978; 3-328.
DR HAWAP; MF_00268; -; 1.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001553; RecA.
DR Pfam; PF00154; RecA; 1.
DR PRINTS; PR00142; RecA.
DR ProDom; PD000229; RecA; 1.
DR SMART; SM00382; AAA; 1.
DR TIGRFAMs; TIGR02012; tigrfam_recA; 1.
DR PROSITE; PS00321; RecA_1; 1.
DR PROSITE; PS0162; RecA_2; 1.
DR PROSITE; PS0163; RecA_3; 1.
KW ATP-binding; Complete proteome; DNA damage; DNA recombination;
KW DNA repair; DNA-binding; Nucleotide-binding; SOS response.
FT INIT MET 0
FT NP_BIND 66 73 ATP (By similarity).
FT SEQUENCE 352 AA; 37813 MW; F24E51CB9A31B818 CRC64;
SQ
Query Match 97.6%; Score 1637; DB 1; Length 352;
Best Local Similarity 97.0%; Pred. No. 4.6e-101;
Matches 325; Conservative 6; Mismatches 4; Indels 0; Gaps 0;
QY 1 AIDENKQALAAALGOIEKQFGKSGIMRLGEDRSMVDVTISTGSLSLDIALGAGGLPMGR 60
Db 1 AIDENKQALAAALGOIEKQFGKSGIMRLGEDRSMVDVTISTGSLSLDIALGAGGLPMGR 60
QY 61 IVEIYGPESSGKTTTLQVIAAAQREGTKTCAFDIAEHALDPTIYARKLGVDIDNLLCSQPD 120
Db 61 IVEIYGPESSGKTTTLQVIAAAQREGTKTCAFDIAEHALDPTIYARKLGVDIDNLLCSQPD 120
QY 121 TGEQALEICDALARSGAVDVIWDSVAALTPKAEIEGIGDSHMGLAARMWSQAMRKLAG 180
Db 121 TGEQALEICDALARSGAVDVIWDSVAALTPKAEIEGIGDSHMGLAARMWSQAMRKLAG 180
QY 181 NLKQNTLLIFINQIRMKIGVMFGNPGTETGGNALKFYASVRLDIRRIGAVKEGNNVGS 240
Db 181 NLKQNTLLIFINQIRMKIGVMFGNPGTETGGNALKFYASVRLDIRRIGAVKEGNNVGS 240
QY 241 ETRVKVKNKTAAPKQAEFQILYEGEINPFYGLVDLGVKEKLIIEKAGWYSYKEKIGQ 300
Db 241 ETRVKVKNKTAAPKQAEFQILYEGEINPFYGLVDLGVKEKLIIEKAGWYSYKEKIGQ 300
QY 301 GKANATAWLNKONPETAKEIEKRVRELLLSNQNPSTP 335
Db 301 GKANATAWLNKONPETAKEIEKRVRELLLSNQNPSTP 335
```

```
Db 301 GKANATTWLNKONPETAKEIEKRVRELLLSNQNPSTP 335
RESULT 7
ID RECA_SALTY STANDARD; PRT; 352 AA.
AC P65977; O8XET0;
DT 28-FEB-2003 (Rel. 41, Created)
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE RecA protein (Recombinase A).
GN Name=recA; OrderedLocusNames=STM2829;
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609; DOI=10.1038/35101614;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
LT2.";
RL Nature 413:852-856 (2001).
CC -!- FUNCTION: Can catalyze the hydrolysis of ATP in the presence of
CC single-stranded DNA, the ATP-dependent uptake of single-stranded
CC DNA by duplex DNA, and the ATP-dependent hybridization of
CC homologous single-stranded DNAs. It interacts with LexA causing
CC its activation and leading to its autocatalytic cleavage (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the recA family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; AE008829; AAL21709.1; -; Genomic_DNA.
DR HSSP; P03017; 2REB.
DR SMR; P65977; 3-328.
DR StyGene; SG???; recA.
DR HAWAP; MF_00268; -; 1.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001553; RecA.
DR Pfam; PF00154; RecA; 1.
DR PRINTS; PR00142; RecA.
DR ProDom; PD000229; RecA; 1.
DR SMART; SM00382; AAA; 1.
DR TIGRFAMs; TIGR02012; tigrfam_recA; 1.
DR PROSITE; PS00321; RecA_1; 1.
DR PROSITE; PS0162; RecA_2; 1.
DR PROSITE; PS0163; RecA_3; 1.
KW ATP-binding; Complete proteome; DNA damage; DNA recombination;
KW DNA repair; DNA-binding; Nucleotide-binding; SOS response.
FT INIT MET 0
FT NP_BIND 66 73 ATP (By similarity).
FT SEQUENCE 352 AA; 37813 MW; F24E51CB9A31B818 CRC64;
SQ
Query Match 97.6%; Score 1637; DB 1; Length 352;
Best Local Similarity 97.0%; Pred. No. 4.6e-101;
Matches 325; Conservative 6; Mismatches 4; Indels 0; Gaps 0;
QY 1 AIDENKQALAAALGOIEKQFGKSGIMRLGEDRSMVDVTISTGSLSLDIALGAGGLPMGR 60
Db 1 AIDENKQALAAALGOIEKQFGKSGIMRLGEDRSMVDVTISTGSLSLDIALGAGGLPMGR 60
QY 61 IVEIYGPESSGKTTTLQVIAAAQREGTKTCAFDIAEHALDPTIYARKLGVDIDNLLCSQPD 120
Db 61 IVEIYGPESSGKTTTLQVIAAAQREGTKTCAFDIAEHALDPTIYARKLGVDIDNLLCSQPD 120
QY 121 TGEQALEICDALARSGAVDVIWDSVAALTPKAEIEGIGDSHMGLAARMWSQAMRKLAG 180
Db 121 TGEQALEICDALARSGAVDVIWDSVAALTPKAEIEGIGDSHMGLAARMWSQAMRKLAG 180
QY 181 NLKQNTLLIFINQIRMKIGVMFGNPGTETGGNALKFYASVRLDIRRIGAVKEGNNVGS 240
Db 181 NLKQNTLLIFINQIRMKIGVMFGNPGTETGGNALKFYASVRLDIRRIGAVKEGNNVGS 240
QY 241 ETRVKVKNKTAAPKQAEFQILYEGEINPFYGLVDLGVKEKLIIEKAGWYSYKEKIGQ 300
Db 241 ETRVKVKNKTAAPKQAEFQILYEGEINPFYGLVDLGVKEKLIIEKAGWYSYKEKIGQ 300
QY 301 GKANATAWLNKONPETAKEIEKRVRELLLSNQNPSTP 335
Db 301 GKANATTWLNKONPETAKEIEKRVRELLLSNQNPSTP 335
```

```

Db      61 IVEIYGPESSGKTTLTQVIAAAQREGKTCAPIDAHALDPVYARKLGVDIDNLLCSQPD 120
QY      121 TGEQALEICDALARSGADVIVVDSVAALTPKAEIEGIGDSHGLAARMMSQAWRKLAG 180
Db      121 TGEQALEICDALARSGADVIVVDSVAALTPKAEIEGIGDSHGLAARMMSQAWRKLAG 180
QY      181 NLKQSNLTLLIFINQIRMKIGVMFGNPNPTTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
Db      181 NLKQSNLTLLIFINQIRMKIGVMFGNPNPTTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
QY      241 ETRVKKVKNKIAAPFKQAEFOILYEGINFGYELVDLGKVKLEKAGAWTSYNGEKIGQ 300
Db      241 ETRVKKVKNKIAAPFKQAEFOILYEGINFGYELVDLGKVKLEKAGAWTSYNGEKIGQ 300
QY      301 GKANATAWLKONPETAKEIEKKVRELLLSNPNSTP 335
Db      301 GKANATTWLKENPATAKEIEKRVRELLLSNQNPATP 335

```

RESULT 8

```

Q57KU4_SALCH
ID Q57KU4_SALCH PRELIMINARY; PRT; 353 AA.
AC Q57KU4;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE DNA strand exchange and recombination protein with protease and
DE nuclease activity.
GN Name=recA; OrderedLocusNames=SC2762;
OS Salmonella choleraesuis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=591;
RN [1]
NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=SC-B67;
RX PubMed=15781495; DOI=10.1093/nar/gki1297;
RA Chiu C.-H., Tang P., Chu C., Hu S., Bao Q., Yu J., Chou Y.-Y.,
RA Wang H.-S., Lee Y.-S.;
RT "The genome sequence of Salmonella enterica serovar Choleraesuis, a
RT highly invasive and resistant zoonotic pathogen.";
RL Nucleic Acids Res. 33:1690-1698(2005).
DR EMBL; AE017220; AAX66668.1; -; Genomic_DNA.
DR SMR; Q57KU4; 4-329.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003684; F:damaged DNA binding; IEA.
DR GO; GO:0008094; F:DNA-dependent ATPase activity; IEA.
DR GO; GO:0017111; F:nucleoside-triphosphatase activity; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR GO; GO:0006281; P:DNA repair; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001553; RecA.
DR Pfam; PF00154; RecA; 1.
DR PRINTS; PR00142; RECA.
DR SMART; SM00382; AAA; 1.
DR TIGRfams; TIGR02012; tigrfam_recA; 1.
DR PROSITE; PS00321; RECA_1; 1.
DR PROSITE; PS0162; RECA_2; 1.
DR PROSITE; PS0163; RECA_3; 1.
DR Complete proteome.
SQ SEQUENCE 353 AA; 37944 MW; ECBBAB042F16362C CRC64;

```

Query Match 97.6%; Score 1637; DB 2; Length 353;
 Best Local Similarity 97.0%; Pred. No. 4.6e-101;
 Matches 325; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

```

QY      1 AIDENKQALAAALGQIEKQFGKGSIMRLGEDRSMVDKTTISGLSLDIALGAGGLPMGR 60
Db      2 AIDENKQALAAALGQIEKQFGKGSIMRLGEDRSMVDKTTISGLSLDIALGAGGLPMGR 61
QY      61 IVEIYGPESSGKTTLTQVIAAAQREGKTCAPIDAHALDPVYARKLGVDIDNLLCSQPD 120

```

```

Db      62 IVEIYGPESSGKTTLTQVIAAAQREGKTCAPIDAHALDPVYARKLGVDIDNLLCSQPD 121
QY      121 TGEQALEICDALARSGADVIVVDSVAALTPKAEIEGIGDSHGLAARMMSQAWRKLAG 180
Db      122 TGEQALEICDALARSGADVIVVDSVAALTPKAEIEGIGDSHGLAARMMSQAWRKLAG 181
QY      181 NLKQSNLTLLIFINQIRMKIGVMFGNPNPTTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
Db      182 NLKQSNLTLLIFINQIRMKIGVMFGNPNPTTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 241
QY      241 ETRVKKVKNKIAAPFKQAEFOILYEGINFGYELVDLGKVKLEKAGAWTSYNGEKIGQ 300
Db      242 ETRVKKVKNKIAAPFKQAEFOILYEGINFGYELVDLGKVKLEKAGAWTSYNGEKIGQ 301
QY      301 GKANATAWLKONPETAKEIEKKVRELLLSNPNSTP 335
Db      302 GKANATTWLKENPATAKEIEKRVRELLLSNQNPATP 336

```

RESULT 9

```

Q5PF15_SALPA
ID Q5PF15_SALPA PRELIMINARY; PRT; 353 AA.
AC Q5PF15;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE RecA protein.
GN Name=recA; OrderedLocusNames=SPA2687;
OS Salmonella paratyphi-a.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=54388;
RN [1]
NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 9150 / SARB42;
RX PubMed=15531882; DOI=10.1038/ng1470;
RA McClelland M., Sanderson K.E., Clifton S.W., Latreille P.,
RA Porwollak S., Sabo A., Meyer R., Bieri T., Ozersky P., McLellan M.,
RA Harkins C.R., Wang C., Nguyen C., Berghoff A., Elliott G.,
RA Leonberg S., Strong C., Du F., Carter J., Kremizki C., Layman D.,
RA Leonard S., Sun H., Fulton L., Nash W., Miner T., Minx P.,
RA Delehaanty K., Fronick C., Magrini V., Nhan M., Warren W., Florea L.,
RA Spieth J., Wilson R.K.;
RT "Comparison of genome degradation in Paratyphi A and Typhi, human-
RT restricted serovars of Salmonella enterica that cause typhoid.";
RL Nat. Genet. 36:1268-1274(2004).
DR EMBL; CP000026; AAV78544.1; -; Genomic_DNA.
DR SMR; Q5PF15; 4-329.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003684; F:damaged DNA binding; IEA.
DR GO; GO:0008094; F:DNA-dependent ATPase activity; IEA.
DR GO; GO:0017111; F:nucleoside-triphosphatase activity; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR GO; GO:0006281; P:DNA repair; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001553; RecA.
DR Pfam; PF00154; RecA; 1.
DR PRINTS; PR00142; RECA.
DR PRODOM; PD000229; RecA; 1.
DR SMART; SM00382; AAA; 1.
DR TIGRfams; TIGR02012; tigrfam_recA; 1.
DR PROSITE; PS00321; RECA_1; 1.
DR PROSITE; PS0162; RECA_2; 1.
DR PROSITE; PS0163; RECA_3; 1.
DR Complete proteome.
SQ SEQUENCE 353 AA; 37962 MW; AD15A0C48E7CF6FD CRC64;

```

Query Match 97.4%; Score 1635; DB 2; Length 353;
 Best Local Similarity 96.7%; Pred. No. 6.2e-101;
 Matches 324; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

```
Qy 1 AIDENKOKALAAALGOIEKQFGKSGIMRLGEDRSMVDKTTISTGSLSDIALGAGGLPMGR 60
Db 2 AIDENKOKALAAALGOIEKQFGKSGIMRLGEDRSMVDKTTISTGSLSDIALGAGGLPMGR 61
Qy 61 IVEIYGPSSGKTTTLTQVIAAAREGKTCFADAEHALDPIYARKLGVDIDNLLCSQPD 120
Dy 62 IVEIYGPSSGKTTTLTQVIAAAREGKTCFADAEHALDPIYARKLGVDIDNLLCSQPD 121
Qy 121 TGEQALICDALARSGAVDVIVDSVAALTPKABIEIGDSHGLAARMMSQAMRKLKAG 180
Dy 122 TGEQALICDALARSGAVDVIVDSVAALTPKABIEIGDSHGLAARMMSQAMRKLKAG 181
Qy 181 NLKOSNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVGS 240
Dy 182 NLKOSNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVGS 241
Qy 241 ETRVKKVKNKIAAPFKQAEFQILYEGEINFGYELVDLGVKEKLEKAGAMYSYKGEKIGQ 300
Dy 242 ETRVKKVKNKIAAPFKQAEFQILYEGEINFGYELVDLGVKEKLEKAGAMYSYKGEKIGQ 301
Qy 301 GKANATAWLKNPETAKEIEKKVRELLLSNPSTP 335
Dy 302 GKANATWLNKENPATAKEIEKRVRELLLSNQATP 336

RESULT 10
RECA PHOLL
ID RECA PHOLL STANDARD; PRT; 355 AA.
AC Q7N7A6;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE RecA protein (Recombinase A).
GN Name=recA; OrderedLocusNames=plu1249;
OS Photorhabdus luminescens (subsp. laumondii).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Photorhabdus.
OX NCBI_TaxID=141679;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=TT01;
RX MEDLINE=22957627; PubMed=14528314; DOI=10.1038/nbt886;
RA Duchaud E., Rusniok C., Frangeul L., Buchrieser C., Givaudan A.,
RA Taouric S., Bocs S., Boursaux-Eude C., Chandler M., Charles J.-F.,
RA Dassa E., Derose S., Derzelle S., Freysinet G., Gaudriault S.,
RA Medigue C., Lanois A., Powell K., Sigulier P., Vincent R., Wingate V.,
RA Zouine M., Glaser P., Boemare N., Danchin A., Kunst F.
RT "The genome sequence of the entomopathogenic bacterium Photorhabdus
RT luminescens".
RL Nat. Biotechnol. 21:1307-1313(2003).
CC -!- FUNCTION: Can catalyze the hydrolysis of ATP in the presence of
CC single-stranded DNA, the ATP-dependent uptake of single-stranded
CC DNA by duplex DNA, and the ATP-dependent hybridization of
CC homologous single-stranded DNAs. It interacts with lexA causing
CC its activation and leading to its autocatalytic cleavage (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the recA family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; BX571863; CAE13543.1; -; Genomic_DNA.
CC SMR; Q7N7A6; 4-329.
CC DR HAMAP; MF_00268; -; 1.
CC DR InterPro; IPR001553; RecA.
CC DR Pfam; PF00154; RecA; 1.
CC DR PRINTS; PR00142; RecA.
CC DR ProDom; PD000229; RecA; 1.
```

```
DR TIGRFAMs; TIGR02012; tigrfam_recA; 1.
DR PROSITE; PS00321; RECA_1; 1.
DR PROSITE; PS00162; RECA_2; 1.
DR PROSITE; PS00163; RECA_3; 1.
KW ATP-binding; Complete proteome; DNA damage; DNA recombination;
KW DNA repair; DNA-binding; Nucleotide-binding; SOS response.
FT NP BIND 67 74 ATP (By similarity).
SQ SEQUENCE 355 AA; 39381 MW; FD659ACED827AEF9 CRC64;

Query Match 92.0%; Score 1543; DB 1; Length 355;
Best Local Similarity 91.8%; Pred. No. 8.5e-95;
Matches 303; Conservative 16; Mismatches 11; Indels 0; Gaps 0;

Qy 1 AIDENKOKALAAALGOIEKQFGKSGIMRLGEDRSMVDKTTISTGSLSDIALGAGGLPMGR 60
Dy 2 AIDENKOKALAAALGOIEKQFGKSGIMRLGEDRSMVDKTTISTGSLSDIALGAGGLPMGR 61
Qy 61 IVEIYGPSSGKTTTLTQVIAAAREGKTCFADAEHALDPIYARKLGVDIDNLLCSQPD 120
Dy 62 IVEIYGPSSGKTTTLTQVIAAAREGKTCFADAEHALDPIYARKLGVDIDNLLCSQPD 121
Qy 121 TGEQALICDALARSGAVDVIVDSVAALTPKABIEIGDSHGLAARMMSQAMRKLKAG 180
Dy 122 TGEQALICDALARSGAVDVIVDSVAALTPKABIEIGDSHGLAARMMSQAMRKLKAG 181
Qy 181 NLKOSNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVGS 240
Dy 182 NLKOSNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGNNVGS 241
Qy 241 ETRVKKVKNKIAAPFKQAEFQILYEGEINFGYELVDLGVKEKLEKAGAMYSYKGEKIGQ 300
Dy 242 ETRVKKVKNKIAAPFKQAEFQILYEGEINFGYELVDLGVKEKLEKAGAMYSYKGEKIGQ 301
Qy 301 GKANATAWLKNPETAKEIEKKVRELLLSN 330
Dy 302 GKANATTYLNKEPEVATELKDCLREMLLN 331

RESULT 11
RECA ERWCT
ID RECA ERWCT STANDARD; PRT; 357 AA.
AC Q6D1S8;
DT 10-MAY-2005 (Rel. 47, Created)
DT 10-MAY-2005 (Rel. 47, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE RecA protein (Recombinase A).
GN Name=recA; OrderedLocusNames=ECA3369;
OS Erwinia carotovora (subsp. atroseptica) (Pectobacterium atrosepticum).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Pectobacterium.
OX NCBI_TaxID=29471;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=SCRI 1043 / ATCC BAA-672;
RX PubMed=15263089; DOI=10.1073/pnas.0402424101;
RA Bell K.S., Sebahia M., Pritchard L., Holden M.T.G., Hyman L.J.,
RA Holvea M.C., Thomson N.R., Bentley S.D., Churcher L.J.C., Mungall K.,
RA Atkin R., Bason N., Brooks K., Chillingworth T., Clark K., Doggett J.,
RA Fraser A., Hance Z., Hauser H., Jagers K., Moule S., Norbertczak H.,
RA Ormond D., Price C., Quail M.A., Sanders M., Walker D., Whitehead S.,
RA Salmund G.P.C., Birch P.R.J., Parkhill J., Toth I.K.
RT "Genome sequence of the enterobacterial phytopathogen Erwinia
RT carotovora subsp. atroseptica and characterization of virulence
RT factors."
RL Proc. Natl. Acad. Sci. U.S.A. 101:11105-11110(2004).
CC -!- FUNCTION: Can catalyze the hydrolysis of ATP in the presence of
CC single-stranded DNA, the ATP-dependent uptake of single-stranded
CC DNA by duplex DNA, and the ATP-dependent hybridization of
CC homologous single-stranded DNAs. It interacts with lexA causing
CC its activation and leading to its autocatalytic cleavage (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the recA family.
```

```

CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; BX50851; CAG76267.1; -; Genomic_DNA.
CC DR O6D1S8; 4-329.
CC DR HAMAP; MF 00268; -; 1.
CC DR InterPro; IPR003593; AAA_ATPase.
CC DR InterPro; IPR001553; RecA.
CC DR Pfam; PF00154; RecA; 1.
CC DR PRINTS; PR00142; RECA.
CC DR ProDom; PD000229; RecA; 1.
CC DR SMART; SM00382; AAA; 1.
CC DR TIGRFAMs; TIGR02012; tigrfam_recA; 1.
CC DR PROSITE; PS00321; RECA_1; 1.
CC DR PROSITE; PS50162; RECA_2; 1.
CC DR PROSITE; PS50163; RECA_3; 1.
CC DR ATP-binding; Complete proteome; DNA damage; DNA recombination;
CC DNA repair; DNA-binding; Nucleotide-binding; SOS response.
CC NP BIND 67 74 ATP (By similarity).
CC SEQUENCE 357 AA; 38284 MW; 49F1FBB82322E0098 CRC64;

Query Match 91.4%; Score 1533; DB 1; Length 357;
Best Local Similarity 91.3%; Pred. No. 4e-94;
Matches 303; Conservative 15; Mismatches 14; Indels 0; Gaps 0;

QY 1 AIDENKOKAALAAALGQIEKQFGKGSIMRLGSDRSMVKTISTGSLSLDIALGAGGLPMGR 60
DB 1 AIDENKOKAALAAALGQIEKQFGKGSIMRLGSDRSMVKTISTGSLSLDIALGAGGLPMGR 61
QY 61 IVEIYGPESSGKTTLTQVIAAAREGKTCAFDAEHALDPIYAKLGVDDINLCSQPD 120
DB 62 IVEIYGPESSGKTTLTQVIAAAREGKTCAFDAEHALDPIYAKLGVDDINLCSQPD 121
QY 121 TGEQALBICDALARSGAVDVIVDSVAALTPKAEIEGIDGSHGLAARWMSQAMRKLKAG 180
DB 122 TGEQALBICDALTSGAVDVIVDSVAALTPKAEIEGIDGSHGLAARWMSQAMRKLKAG 181
QY 181 NLKQSNLLFIINQIRKMGVFNPGNETTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
DB 182 NLKQANTLLFIINQIRKMGVFNPGNETTGGNALKFYASVRLDIRRIGAVKEGENVVGS 241
QY 241 ETRVKKVKNKTAAPFQKAEFQILYGEINFGVGLVDLVGVKELTEKAGAMYSYKGEKIGQ 300
DB 242 ETRVKKVKNKVAAPFQKAEFQILYGEINFGVGLVDLVGVKHLTEKAGAMYSYNGDKIGQ 301
QY 301 GKANATATLWLDKNDPETAKEIEKVKRELLLSNPN 332
DB 302 GKANACNFKLENPTISAELODKLREMLLHKGN 333

RESULT 12
RECA_YERPE STANDARD; PRT; 355 AA.
AC P37858;
DT 01-OCT-1994 (Rel. 30, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE RecA protein (Recombinase A).
GN Name=recA; OrderedLocName=yP03307, y0881, yP0379;
OS Versinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Versinia.
OX NCBI_TaxID=632;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=231;
RC Kryukov V.M., Suchkov I.Y., Sazykin I.S., Mishankin B.N.;
RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.
RN [2]

```

```

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=CO-92 / Biovar Orientalis;
RC MEDLINE=21470413; PubMed=1158636; DOI=10.1038/35097083;
RA Packhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA Prentice M.B., Sebaihia M., James K.D., Churcher C.M., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Tarraga A.-M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M.A., Rutherford K.M.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
RT "Genome sequence of Versinia pestis, the causative agent of plague.";
RL Nature 413:523-527(2001).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=KIM5 / Biovar Mediaevalis;
RC MEDLINE=22137863; PubMed=12142430;
RX DOI=10.1128/JB.184.16.4601-4611.2002;
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.P., Liss P.,
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
RA Perry R.D.;
RT "Genome sequence of Versinia pestis KIM.";
RL J. Bacteriol. 184:4601-4611(2002).
RN [4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=91001 / Biovar Mediaevalis;
RC PubMed=15368893;
RA Song Y., Tong Z., Wang J., Wang L., Guo Z., Han Y., Zhang J., Pei D.,
RA Zhou D., Qin H., Pang X., Han Y., Zhai J., Li M., Cui B., Qi Z.,
RA Jin L., Dai R., Chen F., Li S., Ye C., Du Z., Lin W., Wang J., Yu J.,
RA Yang H., Wang J., Huang P., Yang R.;
RT "Complete genome sequence of Versinia pestis strain 91001, an isolate
RT avirulent to humans.";
RL DNA Res. 11:179-197(2004).
CC -1- FUNCTION: Can catalyze the hydrolysis of ATP in the presence of
CC single-stranded DNA, the ATP-dependent uptake of single-stranded
CC DNA by duplex DNA, and the ATP-dependent hybridization of
CC homologous single-stranded DNAs. It interacts with lexA causing
CC its activation and leading to its autocatalytic cleavage.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the recA family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; X75336; CAA53084.1; -; Genomic DNA.
CC EMBL; AJ414156; CAC92539.1; -; Genomic DNA.
CC EMBL; AE013691; AAM84465.1; -; Genomic DNA.
CC EMBL; AE017128; AAS60652.1; -; Genomic DNA.
CC PIR; AG0401; AG0401.
CC PIR; S37586; S37586.
CC HSP; P03017; 2REB.
CC SMR; P37858; 3-328.
CC HAMAP; MF 00268; -; 1.
CC InterPro; IPR003593; AAA_ATPase.
CC InterPro; IPR001553; RecA.
CC Pfam; PF00154; RecA; 1.
CC PRINTS; PR00142; RECA.
CC ProDom; PD000229; RecA; 1.
CC SMART; SM00382; AAA; 1.
CC TIGRFAMs; TIGR02012; tigrfam_recA; 1.
CC PROSITE; PS00321; RECA_1; 1.
CC PROSITE; PS50162; RECA_2; 1.
CC PROSITE; PS50163; RECA_3; 1.
CC ATP-binding; Complete proteome; DNA damage; DNA recombination;
CC DNA repair; DNA-binding; Nucleotide-binding; SOS response.
CC INIT MET 0 0 By similarity.
CC NP BIND 66 73 ATP (By similarity).
CC CONFLICT 281 281 H -> L (in Ref. 1).

```

```

FT CONFLICT 294 294 N -> Y (in Ref. 1).
FT CONFLICT 315 315 I -> N (in Ref. 1).
SQ SEQUENCE 355 AA; 37755 MW; F800D2D1AD32AB81 CRC64;

Query Match 91.2%; Score 1530; DB 1; Length 355;
Best Local Similarity 91.6%; Pred. No. 6.3e-94;
Matches 304; Conservative 15; Mismatches 13; Indels 0; Gaps 0;

Qy 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVTISTGSLSLDIALGAGGLPMGR 60
Db 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVTISTGSLSLDIALGAGGLPMGR 60

Qy 61 IVEIYGPSSSGKTTTLQVIAAAAREGKTCFAIDAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 61 IVEIYGPSSSGKTTTLQVIAAAAREGKTCFAIDAEHALDPIYARKLGVDIDNLLCSQPD 120

Qy 121 TGEQALEICDALRSGAVDVIIVDSVAALTPKAEIEGEIGDSHMGLAARMWSQAMRKLKAG 180
Db 121 TGEQALEICDALRSGAVDVIIVDSVAALTPKAEIEGEIGDSHMGLAARMWSQAMRKLKAG 180

Qy 181 NLKQNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
Db 181 NLKQNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240

Qy 241 ETRVKVVKNTAAAPFKQAEFQILYEGEINFGELVDLGVEKLEKAGAWSYKGEKIGQ 300
Db 241 ETRVKVVKNTAAAPFKQAEFQILYEGEINFGELVDLGVEKLEKAGAWSYKGEKIGQ 300

Qy 301 GKANATAWLKNPETAKEIEKKVRELLLSNPN 332
Db 301 GKANASNYLKENPAIAAELDKKLEMLNGGN 332

RESULT 13
RECA_YERPS
ID RECA_YERPS STANDARD; PRT; 356 AA.
AC Q66E70;
DT 10-MAY-2005 (Rel. 47, Created)
DT 10-MAY-2005 (Rel. 47, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE RecA protein (Recombinase A).
GN Name=recA; OrderedLocNames=YPTB0823;
OS Yersinia pseudotuberculosis.
OC Bacteriia; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=633;
RN [1]

NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=IP32953 / Serotype I;
RX PubMed=15358858; DOI=10.1073/pnas.0404012101;
RA Chain F.S.G., Carniel E., Larimer F.W., Lamerdin J., Stoutland P.O.,
RA Regala W.M., Georgescu A.M., Vergez L.M., Land M.L., Motin V.L.,
RA Brubaker R.R., Fowler J., Hinnebusch J., Marceau M., Medigue C.,
RA Simonet M., Chenal-Francois V., Souza B., Dacheux D., Elliott J.M.,
RA Derbise A., Hauser L.J., Garcia E.;
RT "Insights into the evolution of Yersinia pestis through whole-genome
RT comparison with Yersinia pseudotuberculosis.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:13826-13831(2004).
CC -!- FUNCTION: Can catalyze the hydrolysis of ATP in the presence of
CC single-stranded DNA, the ATP-dependent uptake of single-stranded
CC DNA by duplex DNA, and the ATP-dependent hybridization of
CC homologous single-stranded DNAs. It interacts with LexA causing
CC its activation and leading to its autocatalytic cleavage (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the recA family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----

```

```

DR EMBL; BX936398; CAH20063.1; -; Genomic_DNA.
DR SMR; Q66E70; 4-329.
DR HAMAP; MF_00268; -; 1.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001553; RecA.
DR Pfam; PF00154; RecA; 1.
DR PRINTS; PR00142; RECA.
DR ProDom; PD000229; RecA; 1.
DR SMART; SM00382; AAA; 1.
DR TIGRFAMs; TIGR02012; tigrfam_recA; 1.
DR PROSITE; PS00321; RECA_1; 1.
DR PROSITE; PS0162; RECA_2; 1.
DR PROSITE; PS0163; RECA_3; 1.
KW ATP-binding; Complete proteome; DNA damage; DNA recombination;
KW DNA repair; DNA-binding; Nucleotide-binding; SOS response.
FT NP_BIND 67 74 ATP (By similarity).
SQ SEQUENCE 356 AA; 37914 MW; DIC0D90F58C8642B CRC64;

Query Match 91.2%; Score 1530; DB 1; Length 356;
Best Local Similarity 91.6%; Pred. No. 6.3e-94;
Matches 304; Conservative 15; Mismatches 13; Indels 0; Gaps 0;

Qy 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVTISTGSLSLDIALGAGGLPMGR 60
Db 2 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVTISTGSLSLDIALGAGGLPMGR 61

Qy 61 IVEIYGPSSSGKTTTLQVIAAAAREGKTCFAIDAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 62 IVEIYGPSSSGKTTTLQVIAAAAREGKTCFAIDAEHALDPIYARKLGVDIDNLLCSQPD 121

Qy 121 TGEQALEICDALRSGAVDVIIVDSVAALTPKAEIEGEIGDSHMGLAARMWSQAMRKLKAG 180
Db 122 TGEQALEICDALRSGAVDVIIVDSVAALTPKAEIEGEIGDSHMGLAARMWSQAMRKLKAG 181

Qy 181 NLKQNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
Db 182 NLKQNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 241

Qy 241 ETRVKVVKNTAAAPFKQAEFQILYEGEINFGELVDLGVEKLEKAGAWSYKGEKIGQ 300
Db 242 ETRVKVVKNTAAAPFKQAEFQILYEGEINFGELVDLGVEKLEKAGAWSYKGEKIGQ 301

Qy 301 GKANATAWLKNPETAKEIEKKVRELLLSNPN 332
Db 302 GKANASNYLKENPAIAAELDKKLEMLNGGN 333

RESULT 14
RECA_ENTAG
ID RECA_ENTAG STANDARD; PRT; 353 AA.
AC P33037;
DT 01-OCT-1993 (Rel. 27, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE RecA protein (Recombinase A).
GN Name=recA;
OS Enterobacter agglomerans (Erwinia herbicola) (Pantoea agglomerans).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Pantoea.
OX NCBI_TaxID=549;
RN [1]
NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RA Rappold C.S.J., Klingmueller W.;
RL Submitted (JAN-1993) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Can catalyze the hydrolysis of ATP in the presence of
CC single-stranded DNA, the ATP-dependent uptake of single-stranded
CC DNA by duplex DNA, and the ATP-dependent hybridization of
CC homologous single-stranded DNAs. It interacts with LexA causing
CC its activation and leading to its autocatalytic cleavage.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the recA family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.

CC EMBL; L03291; AAA91766.1; -; Genomic_DNA.

DR PIR; S31481; S31481.

DR HSSP; P03017; 2REB.

DR SMR; P33037; 3-328.

DR HAMAP; MF 00268; -; 1.

DR InterPro; IPR003593; AAA ATPase.

DR InterPro; IPR001553; RecA.

DR Pfam; PF00154; RecA; 1.

DR PRINTS; PR00142; RECA.

DR ProDom; PD000229; RecA; 1.

DR SMART; SM00382; AAA; 1.

DR TIGRFAMs; TIGR02012; tigrfam_recA; 1.

DR PROSITE; PS00321; RECA_1; 1.

DR PROSITE; PS0162; RECA_2; 1.

DR PROSITE; PS0163; RECA_3; 1.

KW ATP-binding; DNA damage; DNA recombination; DNA repair; DNA-binding;

Nucleotide-binding; SOS response.

FT INIT MET 0 By similarity.

FT NP_BIND 66 73 ATP (By similarity).

SQ SEQUENCE 353 AA; 37767 MW; 960F1F2698CAC46B CRC64;

Query Match 91.1%; Score 1528; DB 1; Length 353;

Best Local Similarity 91.8%; Pred. No. 8.5e-94;

Matches 302; Conservative 16; Mismatches 11; Indels 0; Gaps 0;

Qy 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVKTISTGSLSLDIALGAGGLPMGR 60

Db 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVKTISTGSLSLDIALGAGGLPMGR 60

Qy 61 IVEYIGPSSSGKTTLTQVIAAAREGKTCAFDAEHALDPIYAKLGVDDNLLCSQPD 120

Db 61 IVEYIGPSSSGKTTLTQVIAAAREGKTCAFDAEHALDPIYAKLGVDDNLLCSQPD 120

Qy 121 TGEQALEICDALARGAVDVIVVDSVAALTPKAEIEGIGDSHMGSLAARMWSQAMRKLKAG 180

Db 121 TGEQALEICDALARGAVDVIVVDSVAALTPKAEIEGIGDSHMGSLAARMWSQAMRKLKAG 180

Qy 181 NLKQNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240

Db 181 NLKQNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240

Qy 241 ETRVKVVKNTIAAPFKQAEFQILYGEINFGYELVDLGVKLEKAGAWTSYNGEKIQG 300

Db 241 ETRVKVVKNTIAAPFKQAEFQILYGEINFGYELVDLGVKLEKAGAWTSYNGEKIQG 300

Qy 301 GKANATAWLKDNPETAKEIEKKVRELLLS 329

Db 301 GKANATAYLNKPNKVAALDKKJRDMLLS 329

RESULT 15

RECA_SERMA

ID RECA_SERMA

AC RECA_SERMA

DT 01-AUG-1990 (Rel. 15, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 13-SEP-2005 (Rel. 48, Last annotation update)

DE RecA protein (Recombinase A).

GN Name=recA;

OS Serratia marcescens.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Serratia.

OC NCBI_taxID=615;

RN [1]

RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].

RC STRAIN=SM6;

RX MEDLINE=90094239; PubMed=2152908;

RA Ball T.K., Wasmuth C.R., Braunagel S.C., Benedik M.J.;

RT "Expression of Serratia marcescens extracellular proteins requires
RT recA.";

RL J. Bacteriol. 172:342-349(1990).

CC -1- FUNCTION: Can catalyze the hydrolysis of ATP in the presence of
CC single-stranded DNA, the ATP-dependent uptake of single-stranded
CC DNA by duplex DNA, and the ATP-dependent hybridization of
CC homologous single-stranded DNAs. It interacts with lexA causing
CC its activation and leading to its autocatalytic cleavage.

CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).

CC -1- SIMILARITY: Belongs to the recA family.

CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.

CC EMBL; M22935; AAA28567.1; -; Genomic_DNA.

DR HSSP; P03017; 2REB.

DR SMR; P17479; 3-328.

DR HAMAP; MF 00268; -; 1.

DR InterPro; IPR003593; AAA ATPase.

DR InterPro; IPR001553; RecA.

DR Pfam; PF00154; RecA; 1.

DR PRINTS; PR00142; RECA.

DR ProDom; PD000229; RecA; 1.

DR SMART; SM00382; AAA; 1.

DR TIGRFAMs; TIGR02012; tigrfam_recA; 1.

DR PROSITE; PS00321; RECA_1; 1.

DR PROSITE; PS0162; RECA_2; 1.

DR PROSITE; PS0163; RECA_3; 1.

KW ATP-binding; DNA damage; DNA recombination; DNA repair; DNA-binding;

Nucleotide-binding; SOS response.

FT INIT MET 0 By similarity.

FT NP_BIND 66 73 ATP (By similarity).

SQ SEQUENCE 353 AA; 37778 MW; A97369360970F814 CRC64;

Query Match 90.8%; Score 1523; DB 1; Length 353;

Best Local Similarity 92.1%; Pred. No. 1.8e-93;

Matches 302; Conservative 14; Mismatches 12; Indels 0; Gaps 0;

Qy 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVKTISTGSLSLDIALGAGGLPMGR 60

Db 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVKTISTGSLSLDIALGAGGLPMGR 60

Qy 61 IVEYIGPSSSGKTTLTQVIAAAREGKTCAFDAEHALDPIYAKLGVDDNLLCSQPD 120

Db 61 IVEYIGPSSSGKTTLTQVIAAAREGKTCAFDAEHALDPIYAKLGVDDNLLCSQPD 120

Qy 121 TGEQALEICDALARGAVDVIVVDSVAALTPKAEIEGIGDSHMGSLAARMWSQAMRKLKAG 180

Db 121 TGEQALEICDALARGAVDVIVVDSVAALTPKAEIEGIGDSHMGSLAARMWSQAMRKLKAG 180

Qy 181 NLKQNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240

Db 181 NLKQNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240

Qy 241 ETRVKVVKNTIAAPFKQAEFQILYGEINFGYELVDLGVKLEKAGAWTSYNGEKIQG 300

Db 241 ETRVKVVKNTIAAPFKQAEFQILYGEINFGYELVDLGVKLEKAGAWTSYNGEKIQG 300

Qy 301 GKANATAWLKDNPETAKEIEKKVRELLLS 328

Db 301 GKANATAYLNKPNKVAALDKKJRDMLLS 328

Search completed: February 16, 2006, 01:24:01

Job time : 153 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 16, 2006, 01:24:19 ; Search time 32.5 Seconds
(without alignments)
852.195 Million cell updates/sec

Title: US-10-733-782-3
Perfect score: 1678
Sequence: 1 AIDENKQALAAALGQIEKQ.....AKETEKVRELLSNPNSTP 335

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/1/iaa/5-COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/6-COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/H-COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/pCTUS-COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/RE-COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1674	99.8	353	1	US-08-097-831-6
2	1674	99.8	358	2	US-09-626-410-7
3	1674	99.8	358	2	US-09-626-410-9
4	1674	99.8	358	2	US-09-626-410-14
5	1674	99.8	358	2	US-09-116-188-7
6	1674	99.8	358	2	US-09-116-188-9
7	1674	99.8	358	2	US-09-116-188-14
8	1674	99.8	358	2	US-09-626-047-9
9	1674	99.8	358	2	US-09-626-047-14
10	1674	99.8	358	2	US-09-626-343-7
11	1674	99.8	358	2	US-09-626-343-9
12	1674	99.8	358	2	US-09-626-343-14
13	1674	99.8	358	2	US-09-354-922-8
14	1674	99.8	358	2	US-09-354-922-10
15	1674	99.8	358	2	US-09-354-922-15
16	1674	99.8	358	2	US-09-516-051-7
17	1674	99.8	358	2	US-09-516-051-9
18	1674	99.8	358	2	US-09-516-051-14
19	1674	99.8	358	2	US-09-516-695B-8
20	1674	99.8	358	2	US-09-516-695B-10
21	1674	99.8	358	2	US-09-516-695B-15
22	1674	99.8	358	2	US-09-516-695B-15
23	1669	99.5	358	2	US-09-626-410-8
24	1669	99.5	358	2	US-09-626-410-11
25	1669	99.5	358	2	US-09-116-188-8
26	1669	99.5	358	2	US-09-116-188-11
27	1669	99.5	358	2	US-09-626-047-8

28	1669	99.5	358	2	US-09-626-047-11	Sequence 11, Appl
29	1669	99.5	358	2	US-09-626-343-8	Sequence 8, Appl
30	1669	99.5	358	2	US-09-626-343-11	Sequence 11, Appl
31	1669	99.5	358	2	US-09-354-922-9	Sequence 9, Appl
32	1669	99.5	358	2	US-09-354-922-12	Sequence 12, Appl
33	1669	99.5	358	2	US-09-516-051-8	Sequence 8, Appl
34	1669	99.5	358	2	US-09-516-051-11	Sequence 11, Appl
35	1669	99.5	358	2	US-09-516-695B-9	Sequence 9, Appl
36	1669	99.5	358	2	US-09-516-695B-12	Sequence 12, Appl
37	1656	98.7	358	2	US-09-626-410-12	Sequence 12, Appl
38	1656	98.7	358	2	US-09-116-188-12	Sequence 12, Appl
39	1656	98.7	358	2	US-09-626-047-12	Sequence 12, Appl
40	1656	98.7	358	2	US-09-626-343-12	Sequence 12, Appl
41	1656	98.7	358	2	US-09-354-922-13	Sequence 13, Appl
42	1656	98.7	358	2	US-09-516-051-12	Sequence 12, Appl
43	1656	98.7	358	2	US-09-516-695B-13	Sequence 13, Appl
44	1655	98.6	358	2	US-09-626-410-10	Sequence 10, Appl
45	1655	98.6	358	2	US-09-116-188-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-08-097-831-6
; Sequence 6, Application US/08097831
; Patent No. 5510473
; GENERAL INFORMATION:
; APPLICANT: Camerini-Otero, Rafael D.
; APPLICANT: Angov, Evangelina
; TITLE OF INVENTION: Cloning and Expression of Taq recA
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KNOBBE, MARTENS, OLSON AND BEAR
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: USA
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/097,831
; FILING DATE: 19930726
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Altman, Daniel E.
; REGISTRATION NUMBER: 34,115
; REFERENCE/POCKET NUMBER: NIH066.001A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 714-760-0404
; TELEFAX: 714-760-9502
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 353 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; US-08-097-831-6

Query Match 99.8%; Score 1674; DB 1; Length 353;
Best Local Similarity 99.7%; Pred. No. 3,7e-167;
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AIDENKQALAAALGQIEKQFGKSGMVRGDRSMVKTISTGSLSLDIALGAGGLPMGR 60
|||||

Db 2 AIDENKQKALAAALGOIEKQFGKSGIMRLGEDRSDVETISTGSLSLDIALGAGGLPMGR 61
Qy 61 IVEIYGPSSSGKTTTLQVIAAAQREGKTCAPIDAEHALDPIYARKLGVVDIDNLLCSQPD 120
Db 62 IVEIYGPSSSGKTTTLQVIAAAQREGKTCAPIDAEHALDPIYARKLGVVDIDNLLCSQPD 121
Qy 121 TGEQALEICDALARSGAVDVIVDSVAALTPKAEIEGEIGDSHMGGLAARMMSQAMRKLAG 180
Db 122 TGEQALEICDALARSGAVDVIVDSVAALTPKAEIEGEIGDSHMGGLAARMMSQAMRKLAG 181
Qy 181 NLKQNTLLIFINQIRMKIGVMFGNPTTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
Db 182 NLKQNTLLIFINQIRMKIGVMFGNPTTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 241
Qy 241 ETRVVKVKNKIAAPFKQAEFOILYGEINFGELVDLGKVEKLEKAGAWTSYKGEKIGQ 300
Db 242 ETRVVKVKNKIAAPFKQAEFOILYGEINFGELVDLGKVEKLEKAGAWTSYKGEKIGQ 301
Qy 301 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 335
Db 302 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 336

RESULT 2

US-09-626-410-7
; Sequence 7, Application US/09626410
; Patent No. 6287862
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIA, VENKITSWARAN
; APPLICANT: CASTLE, LINDA
; APPLICANT: KREBER, CLAUD M.
; APPLICANT: BASS, STEVE
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
; FILE REFERENCE: 02-020720US
; CURRENT APPLICATION NUMBER: US/09/626,410
; CURRENT FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 09/116,188
; PRIOR FILING DATE: 07-15-1998
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-626-410-7

Query Match 99.8%; Score 1674; DB 2; Length 358;
Best Local Similarity 99.7%; Pred. No. 3.8e-167;
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AIDENKQKALAAALGOIEKQFGKSGIMRLGEDRSDVETISTGSLSLDIALGAGGLPMGR 60
Db 7 AIDENKQKALAAALGOIEKQFGKSGIMRLGEDRSDVETISTGSLSLDIALGAGGLPMGR 66
Qy 61 IVEIYGPSSSGKTTTLQVIAAAQREGKTCAPIDAEHALDPIYARKLGVVDIDNLLCSQPD 120
Db 67 IVEIYGPSSSGKTTTLQVIAAAQREGKTCAPIDAEHALDPIYARKLGVVDIDNLLCSQPD 126
Qy 121 TGEQALEICDALARSGAVDVIVDSVAALTPKAEIEGEIGDSHMGGLAARMMSQAMRKLAG 180
Db 127 TGEQALEICDALARSGAVDVIVDSVAALTPKAEIEGEIGDSHMGGLAARMMSQAMRKLAG 186
Qy 181 NLKQNTLLIFINQIRMKIGVMFGNPTTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
Db 187 NLKQNTLLIFINQIRMKIGVMFGNPTTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 246

Qy 241 ETRVVKVKNKIAAPFKQAEFOILYGEINFGELVDLGKVEKLEKAGAWTSYKGEKIGQ 300
Db 247 ETRVVKVKNKIAAPFKQAEFOILYGEINFGELVDLGKVEKLEKAGAWTSYKGEKIGQ 306
Qy 301 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 335
Db 307 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 341

RESULT 3

US-09-626-410-9
; Sequence 9, Application US/09626410
; Patent No. 6287862
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIA, VENKITSWARAN
; APPLICANT: CASTLE, LINDA
; APPLICANT: KREBER, CLAUD M.
; APPLICANT: BASS, STEVE
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
; FILE REFERENCE: 02-020720US
; CURRENT APPLICATION NUMBER: US/09/626,410
; CURRENT FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 09/116,188
; PRIOR FILING DATE: 07-15-1998
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-626-410-9

Query Match 99.8%; Score 1674; DB 2; Length 358;
Best Local Similarity 99.7%; Pred. No. 3.8e-167;
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AIDENKQKALAAALGOIEKQFGKSGIMRLGEDRSDVETISTGSLSLDIALGAGGLPMGR 60
Db 7 AIDENKQKALAAALGOIEKQFGKSGIMRLGEDRSDVETISTGSLSLDIALGAGGLPMGR 66
Qy 61 IVEIYGPSSSGKTTTLQVIAAAQREGKTCAPIDAEHALDPIYARKLGVVDIDNLLCSQPD 120
Db 67 IVEIYGPSSSGKTTTLQVIAAAQREGKTCAPIDAEHALDPIYARKLGVVDIDNLLCSQPD 126
Qy 121 TGEQALEICDALARSGAVDVIVDSVAALTPKAEIEGEIGDSHMGGLAARMMSQAMRKLAG 180
Db 127 TGEQALEICDALARSGAVDVIVDSVAALTPKAEIEGEIGDSHMGGLAARMMSQAMRKLAG 186
Qy 181 NLKQNTLLIFINQIRMKIGVMFGNPTTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
Db 187 NLKQNTLLIFINQIRMKIGVMFGNPTTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 246
Qy 241 ETRVVKVKNKIAAPFKQAEFOILYGEINFGELVDLGKVEKLEKAGAWTSYKGEKIGQ 300
Db 247 ETRVVKVKNKIAAPFKQAEFOILYGEINFGELVDLGKVEKLEKAGAWTSYKGEKIGQ 306
Qy 301 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 335
Db 307 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 341

RESULT 4

US-09-626-410-14
; Sequence 14, Application US/09626410
; Patent No. 6287862
; GENERAL INFORMATION:

```
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIA, VENKITSWARAN
; APPLICANT: CASTLE, LINDA
; APPLICANT: KREBBER, CLAUD M.
; APPLICANT: BASS, STEVE
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
; FILE REFERENCE: 02-020720US
; CURRENT APPLICATION NUMBER: US/09/626,410
; CURRENT FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 09/116,188
; PRIOR FILING DATE: 07-15-1998
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: consensus
; OTHER INFORMATION: e. coli sequence
US-09-626-410-14

Query Match          99.8%; Score 1674; DB 2; Length 358;
Best Local Similarity 99.7%; Pred. No. 3.8e-167;
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AIDENKQKALAAALGQIEKQFGKSGIMRLGDRSDMDVKTISTGSLSLDIALGAGGLPMGR 60
Db 7 AIDENKQKALAAALGQIEKQFGKSGIMRLGDRSDMDVKTISTGSLSLDIALGAGGLPMGR 66

Qy 61 IVEIYGPESSGKTTLTLOVIAAAREGKTCAPIDAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 67 IVEIYGPESSGKTTLTLOVIAAAREGKTCAPIDAEHALDPIYARKLGVDIDNLLCSQPD 126

Qy 121 TGEQALEICDALARSGAVDVIVDSVAALTPKAEIEGIGDSHMGSLAARMMSQAMRKLKAG 180
Db 127 TGEQALEICDALARSGAVDVIVDSVAALTPKAEIEGIGDSHMGSLAARMMSQAMRKLKAG 186

Qy 181 NLKQSNLTLLIIFINQIRMKIGVMFGNPETTTGNNALKFYASVRLDIRRIGAVKEGENNVVGS 240
Db 187 NLKQSNLTLLIIFINQIRMKIGVMFGNPETTTGNNALKFYASVRLDIRRIGAVKEGENNVVGS 246

Qy 241 ETRVKVVKNKIAAPFKQAEFQILYGEINFGELVDLGKVKLEKAGAWTSYKGEKIGQ 300
Db 247 ETRVKVVKNKIAAPFKQAEFQILYGEINFGELVDLGKVKLEKAGAWTSYKGEKIGQ 306

Qy 301 GKANATAWKDNPTAKEIEKKVRELLLSNPNSTP 335
Db 307 GKANATAWKDNPTAKEIEKKVRELLLSNPNSTP 341

RESULT 5
US-09-116-188-7
; Sequence 7, Application US/09116188
; Patent No. 6326204
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIA, VENKITSWARAN
; APPLICANT: CASTLE, LINDA
; APPLICANT: KREBBER, CLAUD M.
; APPLICANT: BASS, STEVE
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
; FILE REFERENCE: 02-020720US
; CURRENT APPLICATION NUMBER: US/09/626,410
; CURRENT FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 09/116,188
; PRIOR FILING DATE: 07-15-1998
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-116-188-9

Query Match          99.8%; Score 1674; DB 2; Length 358;
Best Local Similarity 99.7%; Pred. No. 3.8e-167;
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AIDENKQKALAAALGQIEKQFGKSGIMRLGDRSDMDVKTISTGSLSLDIALGAGGLPMGR 60
Db 7 AIDENKQKALAAALGQIEKQFGKSGIMRLGDRSDMDVKTISTGSLSLDIALGAGGLPMGR 66

Qy 61 IVEIYGPESSGKTTLTLOVIAAAREGKTCAPIDAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 67 IVEIYGPESSGKTTLTLOVIAAAREGKTCAPIDAEHALDPIYARKLGVDIDNLLCSQPD 126

Qy 121 TGEQALEICDALARSGAVDVIVDSVAALTPKAEIEGIGDSHMGSLAARMMSQAMRKLKAG 180
Db 127 TGEQALEICDALARSGAVDVIVDSVAALTPKAEIEGIGDSHMGSLAARMMSQAMRKLKAG 186

Qy 181 NLKQSNLTLLIIFINQIRMKIGVMFGNPETTTGNNALKFYASVRLDIRRIGAVKEGENNVVGS 240
Db 187 NLKQSNLTLLIIFINQIRMKIGVMFGNPETTTGNNALKFYASVRLDIRRIGAVKEGENNVVGS 246

Qy 241 ETRVKVVKNKIAAPFKQAEFQILYGEINFGELVDLGKVKLEKAGAWTSYKGEKIGQ 300
Db 247 ETRVKVVKNKIAAPFKQAEFQILYGEINFGELVDLGKVKLEKAGAWTSYKGEKIGQ 306

Qy 301 GKANATAWKDNPTAKEIEKKVRELLLSNPNSTP 335
Db 307 GKANATAWKDNPTAKEIEKKVRELLLSNPNSTP 341

RESULT 6
US-09-116-188-9
; Sequence 9, Application US/09116188
; Patent No. 6326204
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIA, VENKITSWARAN
; APPLICANT: CASTLE, LINDA
; APPLICANT: KREBBER, CLAUD M.
; APPLICANT: BASS, STEVE
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
; FILE REFERENCE: 02-020720US
; CURRENT APPLICATION NUMBER: US/09/116,188
; CURRENT FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-116-188-9

Query Match          99.8%; Score 1674; DB 2; Length 358;
Best Local Similarity 99.7%; Pred. No. 3.8e-167;
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
; TITLE OF INVENTION: SEQUENCE RECOMBINATION
; FILE REFERENCE: 02-020720US
; CURRENT APPLICATION NUMBER: US/09/116,188
; CURRENT FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-116-188-7

Query Match          99.8%; Score 1674; DB 2; Length 358;
Best Local Similarity 99.7%; Pred. No. 3.8e-167;
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AIDENKQKALAAALGQIEKQFGKSGIMRLGDRSDMDVKTISTGSLSLDIALGAGGLPMGR 60
Db 7 AIDENKQKALAAALGQIEKQFGKSGIMRLGDRSDMDVKTISTGSLSLDIALGAGGLPMGR 66

Qy 61 IVEIYGPESSGKTTLTLOVIAAAREGKTCAPIDAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 67 IVEIYGPESSGKTTLTLOVIAAAREGKTCAPIDAEHALDPIYARKLGVDIDNLLCSQPD 126

Qy 121 TGEQALEICDALARSGAVDVIVDSVAALTPKAEIEGIGDSHMGSLAARMMSQAMRKLKAG 180
Db 127 TGEQALEICDALARSGAVDVIVDSVAALTPKAEIEGIGDSHMGSLAARMMSQAMRKLKAG 186

Qy 181 NLKQSNLTLLIIFINQIRMKIGVMFGNPETTTGNNALKFYASVRLDIRRIGAVKEGENNVVGS 240
Db 187 NLKQSNLTLLIIFINQIRMKIGVMFGNPETTTGNNALKFYASVRLDIRRIGAVKEGENNVVGS 246

Qy 241 ETRVKVVKNKIAAPFKQAEFQILYGEINFGELVDLGKVKLEKAGAWTSYKGEKIGQ 300
Db 247 ETRVKVVKNKIAAPFKQAEFQILYGEINFGELVDLGKVKLEKAGAWTSYKGEKIGQ 306

Qy 301 GKANATAWKDNPTAKEIEKKVRELLLSNPNSTP 335
Db 307 GKANATAWKDNPTAKEIEKKVRELLLSNPNSTP 341

RESULT 6
US-09-116-188-9
; Sequence 9, Application US/09116188
; Patent No. 6326204
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIA, VENKITSWARAN
; APPLICANT: CASTLE, LINDA
; APPLICANT: KREBBER, CLAUD M.
; APPLICANT: BASS, STEVE
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
; FILE REFERENCE: 02-020720US
; CURRENT APPLICATION NUMBER: US/09/116,188
; CURRENT FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-116-188-9

Query Match          99.8%; Score 1674; DB 2; Length 358;
Best Local Similarity 99.7%; Pred. No. 3.8e-167;
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVKTISTGSLSLDIALGAGGLPMGR 60
DB 7 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVKTISTGSLSLDIALGAGGLPMGR 66
QY 61 IVEIYGPSSGKTTTLTQVIAAAQREGKTCFADAEHALDPIYARKLGVDIDNLLCSQPD 120
DB 67 IVEIYGPSSGKTTTLTQVIAAAQREGKTCFADAEHALDPIYARKLGVDIDNLLCSQPD 126
QY 121 TGEQALEICDALARGAVDVIIVDSVAALTPKAEIEGEIGDGHMGLAARMMSQAMRKLAG 180
DB 127 TGEQALEICDALARGAVDVIIVDSVAALTPKAEIEGEIGDGHMGLAARMMSQAMRKLAG 186
QY 181 NLKQNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
DB 187 NLKQNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 246
QY 241 ETRVVKVKNKIAAPPKQAEFQILYGEINFGELVDLGKVEKLI EKAGAWYSYKGEKIGQ 300
DB 247 ETRVVKVKNKIAAPPKQAEFQILYGEINFGELVDLGKVEKLI EKAGAWYSYKGEKIGQ 306
QY 301 GKANATAWLKONPETAKIEKKVRELLLSNPNSTP 335
DB 307 GKANATAWLKONPETAKIEKKVRELLLSNPNSTP 341

RESULT 7

US-09-116-188-14
; Sequence 14, Application US/09116188
; Patent No. 6326204
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P. C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIA, VENKITSWARAN
; APPLICANT: CASTLE, LINDA
; APPLICANT: KREBBER, CLAUD M.
; APPLICANT: BASS, STEVE
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
; FILE REFERENCE: 02-020720US
; CURRENT APPLICATION NUMBER: US/09/116,188
; CURRENT FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: consensus
US-09-116-188-14

Query Match 99.8%; Score 1674; DB 2; Length 358;
Best Local Similarity 99.7%; Pred. No. 3.8e-167;
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVKTISTGSLSLDIALGAGGLPMGR 60
DB 7 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVKTISTGSLSLDIALGAGGLPMGR 66
QY 61 IVEIYGPSSGKTTTLTQVIAAAQREGKTCFADAEHALDPIYARKLGVDIDNLLCSQPD 120
DB 67 IVEIYGPSSGKTTTLTQVIAAAQREGKTCFADAEHALDPIYARKLGVDIDNLLCSQPD 126
QY 121 TGEQALEICDALARGAVDVIIVDSVAALTPKAEIEGEIGDGHMGLAARMMSQAMRKLAG 180
DB 127 TGEQALEICDALARGAVDVIIVDSVAALTPKAEIEGEIGDGHMGLAARMMSQAMRKLAG 186
QY 181 NLKQNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240

DB 187 NLKQNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 246
QY 241 ETRVVKVKNKIAAPPKQAEFQILYGEINFGELVDLGKVEKLI EKAGAWYSYKGEKIGQ 300
DB 247 ETRVVKVKNKIAAPPKQAEFQILYGEINFGELVDLGKVEKLI EKAGAWYSYKGEKIGQ 306
QY 301 GKANATAWLKONPETAKIEKKVRELLLSNPNSTP 335
DB 307 GKANATAWLKONPETAKIEKKVRELLLSNPNSTP 341

RESULT 8

US-09-626-047-7
; Sequence 7, Application US/09626047
; Patent No. 6335198
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P. C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIA, VENKITSWARAN
; APPLICANT: CASTLE, LINDA
; APPLICANT: KREBBER, CLAUD M.
; APPLICANT: BASS, STEVE
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
; FILE REFERENCE: 02-020720US
; CURRENT APPLICATION NUMBER: US/09/626,047
; CURRENT FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 09/116,188
; PRIOR FILING DATE: 07-15-1998
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-626-047-7

Query Match 99.8%; Score 1674; DB 2; Length 358;
Best Local Similarity 99.7%; Pred. No. 3.8e-167;
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVKTISTGSLSLDIALGAGGLPMGR 60
DB 7 AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMVDVKTISTGSLSLDIALGAGGLPMGR 66
QY 61 IVEIYGPSSGKTTTLTQVIAAAQREGKTCFADAEHALDPIYARKLGVDIDNLLCSQPD 120
DB 67 IVEIYGPSSGKTTTLTQVIAAAQREGKTCFADAEHALDPIYARKLGVDIDNLLCSQPD 126
QY 121 TGEQALEICDALARGAVDVIIVDSVAALTPKAEIEGEIGDGHMGLAARMMSQAMRKLAG 180
DB 127 TGEQALEICDALARGAVDVIIVDSVAALTPKAEIEGEIGDGHMGLAARMMSQAMRKLAG 186
QY 181 NLKQNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
DB 187 NLKQNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 246
QY 241 ETRVVKVKNKIAAPPKQAEFQILYGEINFGELVDLGKVEKLI EKAGAWYSYKGEKIGQ 300
DB 247 ETRVVKVKNKIAAPPKQAEFQILYGEINFGELVDLGKVEKLI EKAGAWYSYKGEKIGQ 306
QY 301 GKANATAWLKONPETAKIEKKVRELLLSNPNSTP 335
DB 307 GKANATAWLKONPETAKIEKKVRELLLSNPNSTP 341

RESULT 9

US-09-626-047-9

```
; Sequence 9, Application US/09626047
; Patent No. 6335198
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIA, VENKITSWARAN
; APPLICANT: CASTLE, LINDA
; APPLICANT: KREBER, CLAUD M.
; APPLICANT: BASS, STEVE
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
; FILE REFERENCE: 02-020720US
; CURRENT APPLICATION NUMBER: US/09/626,047
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 09/116,188
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Escherichia coli
; US-09-626-047-9

Query Match      99.8%; Score 1674; DB 2; Length 358;
Best Local Similarity 99.7%; Pred. No. 3.8e-167;
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AIDENKOKALAAALGQIEKQFGKGSIMRLGDRSDMDVKTISTGSLDIALGAGGLPMGR 60
Db 7 AIDENKOKALAAALGQIEKQFGKGSIMRLGDRSDMDVKTISTGSLDIALGAGGLPMGR 66

Qy 61 IVEIYGPESSGKTTLTQVIAAAREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 67 IVEIYGPESSGKTTLTQVIAAAREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 126

Qy 121 TGEQALETCDALARSAGVDVIVDSVAALTPKAEIEGISHMGLAARMMSQAMRKLKAG 180
Db 127 TGEQALETCDALARSAGVDVIVDSVAALTPKAEIEGISHMGLAARMMSQAMRKLKAG 186

Qy 181 NLKQSTLLIFINQIRMKIGVMFGNPTTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
Db 187 NLKQSTLLIFINQIRMKIGVMFGNPTTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 246

Qy 241 ETRVKKVKNKIAAPFKQAEFOILYGEINFGYELVDLGVKEKLIKAGAWTSYKGEKIQ 300
Db 247 ETRVKKVKNKIAAPFKQAEFOILYGEINFGYELVDLGVKEKLIKAGAWTSYKGEKIQ 306

Qy 301 GKANATAWKONPETAKEIEKKVRELLLSNPSTP 335
Db 307 GKANATAWKONPETAKEIEKKVRELLLSNPSTP 341

RESULT 10
US-09-626-047-14
; Sequence 14, Application US/09626047
; Patent No. 6335198
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIA, VENKITSWARAN
; APPLICANT: CASTLE, LINDA
; APPLICANT: KREBER, CLAUD M.
; APPLICANT: BASS, STEVE
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
```

```
; TITLE OF INVENTION: SEQUENCE RECOMBINATION
; FILE REFERENCE: 02-020720US
; CURRENT APPLICATION NUMBER: US/09/626,047
; CURRENT FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 09/116,188
; PRIOR FILING DATE: 07-15-1998
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: consensus
; OTHER INFORMATION: e. coli sequence
; US-09-626-047-14

Query Match      99.8%; Score 1674; DB 2; Length 358;
Best Local Similarity 99.7%; Pred. No. 3.8e-167;
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AIDENKOKALAAALGQIEKQFGKGSIMRLGDRSDMDVKTISTGSLDIALGAGGLPMGR 60
Db 7 AIDENKOKALAAALGQIEKQFGKGSIMRLGDRSDMDVKTISTGSLDIALGAGGLPMGR 66

Qy 61 IVEIYGPESSGKTTLTQVIAAAREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 67 IVEIYGPESSGKTTLTQVIAAAREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 126

Qy 121 TGEQALETCDALARSAGVDVIVDSVAALTPKAEIEGISHMGLAARMMSQAMRKLKAG 180
Db 127 TGEQALETCDALARSAGVDVIVDSVAALTPKAEIEGISHMGLAARMMSQAMRKLKAG 186

Qy 181 NLKQSTLLIFINQIRMKIGVMFGNPTTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
Db 187 NLKQSTLLIFINQIRMKIGVMFGNPTTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 246

Qy 241 ETRVKKVKNKIAAPFKQAEFOILYGEINFGYELVDLGVKEKLIKAGAWTSYKGEKIQ 300
Db 247 ETRVKKVKNKIAAPFKQAEFOILYGEINFGYELVDLGVKEKLIKAGAWTSYKGEKIQ 306

Qy 301 GKANATAWKONPETAKEIEKKVRELLLSNPSTP 335
Db 307 GKANATAWKONPETAKEIEKKVRELLLSNPSTP 341

RESULT 11
US-09-626-343-7
; Sequence 7, Application US/09626343
; Patent No. 6352859
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW P.C.
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIA, VENKITSWARAN
; APPLICANT: CASTLE, LINDA
; APPLICANT: KREBER, CLAUD M.
; APPLICANT: BASS, STEVE
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
; FILE REFERENCE: 02-020720US
; CURRENT APPLICATION NUMBER: US/09/626,343
; CURRENT FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 09/116,188
; PRIOR FILING DATE: 07-15-1998
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 358
; TYPE: PRT
```

```
; ORGANISM: Escherichia coli
US-09-626-343-7

Query Match          99.8%; Score 1674; DB 2; Length 358;
Best Local Similarity 99.7%; Pred. No. 3.8e-167;
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AIDENKOKALAAALGQIEKQFGKGSIMRLGEDRSDMDVKTISTGSLSLDIALGAGGLPMGR 60
Db 7 AIDENKOKALAAALGQIEKQFGKGSIMRLGEDRSDMDVKTISTGSLSLDIALGAGGLPMGR 66
Qy 61 IVEIYGPESSGKTTTLTQVIAAAQREGKTCFADAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 67 IVEIYGPESSGKTTTLTQVIAAAQREGKTCFADAEHALDPIYARKLGVDIDNLLCSQPD 126
Qy 121 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGIGDSHMGLAARMMSQAMRKLAG 180
Db 127 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGIGDSHMGLAARMMSQAMRKLAG 186
Qy 181 NLKQNTLLIFINQIRMKIGVMFGNPEPTTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
Db 187 NLKQNTLLIFINQIRMKIGVMFGNPEPTTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 246
Qy 241 ETRVWVKNKIAAPFKQAEFOILYGEGINFYGELVDLGVKEKLEKAGAWTSYKGEKIGQ 300
Db 247 ETRVWVKNKIAAPFKQAEFOILYGEGINFYGELVDLGVKEKLEKAGAWTSYKGEKIGQ 306
Qy 301 GKANATAWLKONPETAKIEKKVRELLLSNPSTP 335
Db 307 GKANATAWLKONPETAKIEKKVRELLLSNPSTP 341

RESULT 12
US-09-626-343-9
; Sequence 9, Application US/09626343
; Patent No. 6352859
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P. C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIA, VENKITSWARAN
; APPLICANT: CASTLE, LINDA
; APPLICANT: KREBER, CLAUD M.
; APPLICANT: BASS, STEVE
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
; FILE REFERENCE: 02-020727US
; CURRENT APPLICATION NUMBER: US/09/626,343
; CURRENT FILING DATE: 2000-07-26
; PRIOR FILING DATE: 07-15-1998
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Escherichia coli
; OTHER INFORMATION: Description of Artificial Sequence: consensus
; OTHER INFORMATION: e. coli sequence
US-09-626-343-9

Query Match          99.8%; Score 1674; DB 2; Length 358;
Best Local Similarity 99.7%; Pred. No. 3.8e-167;
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AIDENKOKALAAALGQIEKQFGKGSIMRLGEDRSDMDVKTISTGSLSLDIALGAGGLPMGR 60
Db 7 AIDENKOKALAAALGQIEKQFGKGSIMRLGEDRSDMDVKTISTGSLSLDIALGAGGLPMGR 66
Qy 61 IVEIYGPESSGKTTTLTQVIAAAQREGKTCFADAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 67 IVEIYGPESSGKTTTLTQVIAAAQREGKTCFADAEHALDPIYARKLGVDIDNLLCSQPD 126
Qy 121 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGIGDSHMGLAARMMSQAMRKLAG 180
Db 127 TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGIGDSHMGLAARMMSQAMRKLAG 186
Qy 181 NLKQNTLLIFINQIRMKIGVMFGNPEPTTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
Db 187 NLKQNTLLIFINQIRMKIGVMFGNPEPTTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 246
Qy 241 ETRVWVKNKIAAPFKQAEFOILYGEGINFYGELVDLGVKEKLEKAGAWTSYKGEKIGQ 300
Db 247 ETRVWVKNKIAAPFKQAEFOILYGEGINFYGELVDLGVKEKLEKAGAWTSYKGEKIGQ 306

; ORGANISM: Escherichia coli
US-09-626-343-9

Query Match          99.8%; Score 1674; DB 2; Length 358;
Best Local Similarity 99.7%; Pred. No. 3.8e-167;
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AIDENKOKALAAALGQIEKQFGKGSIMRLGEDRSDMDVKTISTGSLSLDIALGAGGLPMGR 60
Db 7 AIDENKOKALAAALGQIEKQFGKGSIMRLGEDRSDMDVKTISTGSLSLDIALGAGGLPMGR 66
Qy 61 IVEIYGPESSGKTTTLTQVIAAAQREGKTCFADAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 67 IVEIYGPESSGKTTTLTQVIAAAQREGKTCFADAEHALDPIYARKLGVDIDNLLCSQPD 126
```

Db 247 ETRVVKVKNKIAAPFKQAEFOILYGEINFGYELVDLGVKEKLEKAGAWTSYKGEKIGQ 306
Qy 301 GKANATAWLKNPETAKEIEKKVRELLSNPNSTP 335
Db 307 GKANATAWLKNPETAKEIEKKVRELLSNPNSTP 341

RESULT 14

US-09-354-922-8
; Sequence 8, Application US/09354922
; Patent No. 6379964

GENERAL INFORMATION:

; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIAN, VENKITESWATAN
; APPLICANT: CASTLE, LINDA A.
; APPLICANT: KREBER, CLAUD M.
; APPLICANT: BASS, STEVE
; APPLICANT: ZHANG, YING-XIN
; APPLICANT: COX, TONY
; APPLICANT: HUISMAN, GJALT
; APPLICANT: YUAN, LING
; APPLICANT: AFFHOLTER, JOSEPH A.

; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
; FILE REFERENCE: 02-020730US
; CURRENT APPLICATION NUMBER: US/09/354,922
; CURRENT FILING DATE: 1999-07-15
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 8

; LENGTH: 358

; TYPE: PRT

; ORGANISM: Escherichia coli

US-09-354-922-8

Query Match 99.8%; Score 1674; DB 2; Length 358;

Best Local Similarity 99.7%; Pred. No. 3.8e-167;

Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AIDENKOKALAAALGOIEKQFGKGSIMRLGDRSDMDVKTISTGSLDIALGAGGLPMGR 60
Db 7 AIDENKOKALAAALGOIEKQFGKGSIMRLGDRSDMDVKTISTGSLDIALGAGGLPMGR 66

Qy 61 IVEIYGPSSGKTTTLTQVIAAAQREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 67 IVEIYGPSSGKTTTLTQVIAAAQREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 126

Qy 121 TGEQALIECDALARSQAVDVIVDSVAALTPKABIEGIEGDSHGLAARMMSQAMRKLKAG 180
Db 127 TGEQALIECDALARSQAVDVIVDSVAALTPKABIEGIEGDSHGLAARMMSQAMRKLKAG 186

Qy 181 NLKQSNLLIFINQIRMKIGVMFCNPETTTGGNALKFYASVRLDIRRIGAVKEGENVYGS 240
Db 187 NLKQSNLLIFINQIRMKIGVMFCNPETTTGGNALKFYASVRLDIRRIGAVKEGENVYGS 246

Qy 241 ETRVVKVKNKIAAPFKQAEFOILYGEINFGYELVDLGVKEKLEKAGAWTSYKGEKIGQ 300
Db 247 ETRVVKVKNKIAAPFKQAEFOILYGEINFGYELVDLGVKEKLEKAGAWTSYKGEKIGQ 306

Qy 301 GKANATAWLKNPETAKEIEKKVRELLSNPNSTP 335

Db 307 GKANATAWLKNPETAKEIEKKVRELLSNPNSTP 341

RESULT 15

US-09-354-922-10

; Sequence 10, Application US/09354922

; Patent No. 6379964

GENERAL INFORMATION:

; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIAN, VENKITESWATAN
; APPLICANT: CASTLE, LINDA A.
; APPLICANT: KREBER, CLAUD M.
; APPLICANT: BASS, STEVE
; APPLICANT: ZHANG, YING-XIN
; APPLICANT: COX, TONY
; APPLICANT: HUISMAN, GJALT
; APPLICANT: YUAN, LING
; APPLICANT: AFFHOLTER, JOSEPH A.

; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
; FILE REFERENCE: 02-020730US
; CURRENT APPLICATION NUMBER: US/09/354,922
; CURRENT FILING DATE: 1999-07-15
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 10

; LENGTH: 358

; TYPE: PRT

; ORGANISM: Escherichia coli

US-09-354-922-10

Query Match 99.8%; Score 1674; DB 2; Length 358;

Best Local Similarity 99.7%; Pred. No. 3.8e-167;

Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AIDENKOKALAAALGOIEKQFGKGSIMRLGDRSDMDVKTISTGSLDIALGAGGLPMGR 60
Db 7 AIDENKOKALAAALGOIEKQFGKGSIMRLGDRSDMDVKTISTGSLDIALGAGGLPMGR 66

Qy 61 IVEIYGPSSGKTTTLTQVIAAAQREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 67 IVEIYGPSSGKTTTLTQVIAAAQREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 126

Qy 121 TGEQALIECDALARSQAVDVIVDSVAALTPKABIEGIEGDSHGLAARMMSQAMRKLKAG 180
Db 127 TGEQALIECDALARSQAVDVIVDSVAALTPKABIEGIEGDSHGLAARMMSQAMRKLKAG 186

Qy 181 NLKQSNLLIFINQIRMKIGVMFCNPETTTGGNALKFYASVRLDIRRIGAVKEGENVYGS 240
Db 187 NLKQSNLLIFINQIRMKIGVMFCNPETTTGGNALKFYASVRLDIRRIGAVKEGENVYGS 246

Qy 241 ETRVVKVKNKIAAPFKQAEFOILYGEINFGYELVDLGVKEKLEKAGAWTSYKGEKIGQ 300
Db 247 ETRVVKVKNKIAAPFKQAEFOILYGEINFGYELVDLGVKEKLEKAGAWTSYKGEKIGQ 306

Qy 301 GKANATAWLKNPETAKEIEKKVRELLSNPNSTP 335

Db 307 GKANATAWLKNPETAKEIEKKVRELLSNPNSTP 341

Search completed: February 16, 2006, 01:26:08

Job time : 32.5 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: February 16, 2006, 01:39:50 ; Search time 123 Seconds
(without alignments)
1137.990 Million cell updates/sec

Title: US-10-733-782-3

Perfect score: 1678

Sequence: 1 AIDENKQKALAAALGQIEKQ.....AKEIEKKVRELLSNPNSTP 335

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Published Applications_AA_Main:*
- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
 - 2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
 - 3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
 - 4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
 - 5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
 - 6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1678	100.0	335	4	US-10-733-782-3
2	1674	99.8	335	4	US-10-733-782-1
3	1674	99.8	352	4	US-10-733-782-5
4	1674	99.8	353	4	US-10-282-122A-43319
5	1674	99.8	358	4	US-10-194-686-8
6	1674	99.8	358	4	US-10-194-686-10
7	1674	99.8	358	4	US-10-194-686-15
8	1669	99.5	358	4	US-10-194-686-9
9	1669	99.5	358	4	US-10-194-686-12
10	1656	98.7	358	4	US-10-194-686-13
11	1655	98.6	358	4	US-10-194-686-11
12	1637	97.6	353	4	US-10-282-122A-75948
13	1609	95.9	352	4	US-10-282-122A-59525
14	1589	94.7	352	4	US-10-282-122A-55632
15	1530	91.2	356	4	US-10-282-122A-78524
16	1500	89.4	355	4	US-10-282-122A-68593
17	1463	87.2	351	4	US-10-282-122A-72895
18	1432	85.3	412	4	US-10-282-122A-77075
19	1349	80.4	354	4	US-10-282-122A-67415
20	1328	79.1	354	4	US-10-282-122A-58235
21	1268	75.6	346	4	US-10-282-122A-66505
22	1257	74.9	355	4	US-10-282-122A-67942
23	1255	74.8	349	4	US-10-282-122A-44768
24	1255	74.8	369	4	US-10-282-122A-69802
25	1242.5	74.0	349	4	US-10-282-122A-63392
26	1238.5	73.8	342	4	US-10-751-928-20
27	1237.5	73.7	348	4	US-10-282-122A-61049

28	1213.5	72.3	348	5	US-10-988-943-33	Sequence 33, Appl
29	1212.5	72.3	348	4	US-10-282-122A-65975	Sequence 65975, A
30	1208.5	72.0	375	4	US-10-282-122A-65031	Sequence 65031, A
31	1200.5	71.5	356	4	US-10-282-122A-47848	Sequence 47848, A
32	1200.5	71.5	356	4	US-10-282-122A-50519	Sequence 50519, A
33	1196.5	71.3	356	4	US-10-282-122A-49452	Sequence 49452, A
34	1180.5	70.4	353	4	US-10-282-122A-51342	Sequence 51342, A
35	1148	68.4	377	4	US-10-156-761-10029	Sequence 10029, A
36	1139	67.9	356	4	US-10-282-122A-52244	Sequence 52244, A
37	1135	67.6	354	4	US-10-282-122A-53057	Sequence 53057, A
38	1131	67.4	349	4	US-10-282-122A-51953	Sequence 51953, A
39	1112	66.3	350	4	US-10-282-122A-61898	Sequence 61898, A
40	1107	66.0	350	4	US-10-203-927A-2	Sequence 2, Appli
41	1107	66.0	356	5	US-10-501-282-5628	Sequence 5628, Ap
42	1107	66.0	361	5	US-10-501-282-5630	Sequence 5630, Ap
43	1095	65.3	348	4	US-10-282-122A-71634	Sequence 71634, A
44	1089	64.9	347	4	US-10-282-122A-70318	Sequence 70318, A
45	1089	64.9	357	4	US-10-724-972A-4382	Sequence 4382, Ap

ALIGNMENTS

RESULT 1
US-10-733-782-3
; Sequence 3, Application US/10733782
; Publication No. US20040157248A1
; GENERAL INFORMATION:
; APPLICANT: Cox, Michael
; APPLICANT: Lusetti, Shelley
; APPLICANT: Eggleter, Aimee
; TITLE OF INVENTION: RecA Mutants
; FILE REFERENCE: 960296.99501
; CURRENT APPLICATION NUMBER: US/10733,782
; CURRENT FILING DATE: 2003-12-11
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 3
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-733-782-3

Query Match		100.0%;	Score 1678;	DB 4;	Length 335;
Best Local Similarity		100.0%;	Pred. No. 2.7e-152;		
Matches 335;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	AIDENKQKALAAALGQIEKQFGKGSIMRLGDRSDMVKTISTGSLSLDIALGAGGLPMGR	60		
Db	1	AIDENKQKALAAALGQIEKQFGKGSIMRLGDRSDMVKTISTGSLSLDIALGAGGLPMGR	60		
Qy	61	IVEIYGPSSGKTTTLQVIAAAREGKTCAPIDAHAALDPIYARKLGVDIDNLLCSQPD	120		
Db	61	IVEIYGPSSGKTTTLQVIAAAREGKTCAPIDAHAALDPIYARKLGVDIDNLLCSQPD	120		
Qy	121	TGEALRICDALARSGAVDVIVDSVAALTPKAEIEGIGDSHMGSLAARMMQAMRKLKAG	180		
Db	121	TGEALRICDALARSGAVDVIVDSVAALTPKAEIEGIGDSHMGSLAARMMQAMRKLKAG	180		
Qy	181	NLKQSNLLIFINQIRMKIGVMFGNPTTTGGNNALKFYASVRLDIRRIGAVKEGNNVGS	240		
Db	181	NLKQSNLLIFINQIRMKIGVMFGNPTTTGGNNALKFYASVRLDIRRIGAVKEGNNVGS	240		
Qy	241	ETRVKVVKNKIAAPFKQAEFQILYGEINPFYGVLDLGVKEKLEKAGAWYSYKGEKIQ	300		
Db	241	ETRVKVVKNKIAAPFKQAEFQILYGEINPFYGVLDLGVKEKLEKAGAWYSYKGEKIQ	300		
Qy	301	GKANATAWLNDPNPTAKEIEKKVRELLSNPNSTP	335		
Db	301	GKANATAWLNDPNPTAKEIEKKVRELLSNPNSTP	335		

RESULT 2

US-10-733-782-1
; Sequence 1, Application US/10733782
; Publication No. US20040157248A1
; GENERAL INFORMATION:
; APPLICANT: Cox, Michael
; APPLICANT: Lusetti, Shelley
; APPLICANT: Egger, Aimee
; TITLE OF INVENTION: RecA Mutants
; FILE REFERENCE: 960296.99501
; CURRENT APPLICATION NUMBER: US/10/733,782
; CURRENT FILING DATE: 2003-12-11
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-733-782-1

Query Match 99.8%; Score 1674; DB 4; Length 335;
Best Local Similarity 99.7%; Pred. No. 6.5e-152;
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy	1	AIDENKQKALAAALGOIEKQFGKGSIMRLGEDRSDMDVETISTGSLSDIALGAGGLPMGR	60
Db	1	AIDENKQKALAAALGOIEKQFGKGSIMRLGEDRSDMDVETISTGSLSDIALGAGGLPMGR	60
Qy	61	IVEIYGPESSGKTTTLTQVIAAAQREGKTCFADAEHALDPIYARKLGVDIDNLLCSQPD	120
Db	61	IVEIYGPESSGKTTTLTQVIAAAQREGKTCFADAEHALDPIYARKLGVDIDNLLCSQPD	120
Qy	121	TGQALEICDALARSGAVDIVVDSVAALTPKAEIEGEIGDSDHMGLAARMMSQAMRKLAG	180
Db	121	TGQALEICDALARSGAVDIVVDSVAALTPKAEIEGEIGDSDHMGLAARMMSQAMRKLAG	180
Qy	181	NLKQSNLTLLIFINQIRMKIGVMFGNPEPTTGGNALKFYASVRLDIRRIGAVKEGENVVG	240
Db	181	NLKQSNLTLLIFINQIRMKIGVMFGNPEPTTGGNALKFYASVRLDIRRIGAVKEGENVVG	240
Qy	241	ETRVKVVKNKIAAPFKQAEFQILYGBGINFYGBELVDLGVKEKLEKAGAWYSYKGEKIQ	300
Db	241	ETRVKVVKNKIAAPFKQAEFQILYGBGINFYGBELVDLGVKEKLEKAGAWYSYKGEKIQ	300
Qy	301	GKANATAWLKNDPETAKEIEKKVRELLSNPNSTP	335
Db	301	GKANATAWLKNDPETAKEIEKKVRELLSNPNSTP	335

RESULT 3
US-10-733-782-5
; Sequence 5, Application US/10733782
; Publication No. US20040157248A1
; GENERAL INFORMATION:
; APPLICANT: Cox, Michael
; APPLICANT: Lusetti, Shelley
; APPLICANT: Egger, Aimee
; TITLE OF INVENTION: RecA Mutants
; FILE REFERENCE: 960296.99501
; CURRENT APPLICATION NUMBER: US/10/733,782
; CURRENT FILING DATE: 2003-12-11
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-733-782-5

Query Match 99.8%; Score 1674; DB 4; Length 352;
Best Local Similarity 99.7%; Pred. No. 7e-152;
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy	1	AIDENKQKALAAALGOIEKQFGKGSIMRLGEDRSDMDVETISTGSLSDIALGAGGLPMGR	60
----	---	--	----

US-10-733-782-1

US-10-733-782-1
; Sequence 1, Application US/10733782
; Publication No. US20040157248A1
; GENERAL INFORMATION:
; APPLICANT: Cox, Michael
; APPLICANT: Lusetti, Shelley
; APPLICANT: Egger, Aimee
; TITLE OF INVENTION: RecA Mutants
; FILE REFERENCE: 960296.99501
; CURRENT APPLICATION NUMBER: US/10/733,782
; CURRENT FILING DATE: 2003-12-11
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-733-782-1

Query Match 99.8%; Score 1674; DB 4; Length 335;
Best Local Similarity 99.7%; Pred. No. 6.5e-152;
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy	1	AIDENKQKALAAALGOIEKQFGKGSIMRLGEDRSDMDVETISTGSLSDIALGAGGLPMGR	60
Db	1	AIDENKQKALAAALGOIEKQFGKGSIMRLGEDRSDMDVETISTGSLSDIALGAGGLPMGR	60
Qy	61	IVEIYGPESSGKTTTLTQVIAAAQREGKTCFADAEHALDPIYARKLGVDIDNLLCSQPD	120
Db	61	IVEIYGPESSGKTTTLTQVIAAAQREGKTCFADAEHALDPIYARKLGVDIDNLLCSQPD	120
Qy	121	TGQALEICDALARSGAVDIVVDSVAALTPKAEIEGEIGDSDHMGLAARMMSQAMRKLAG	180
Db	121	TGQALEICDALARSGAVDIVVDSVAALTPKAEIEGEIGDSDHMGLAARMMSQAMRKLAG	180
Qy	181	NLKQSNLTLLIFINQIRMKIGVMFGNPEPTTGGNALKFYASVRLDIRRIGAVKEGENVVG	240
Db	181	NLKQSNLTLLIFINQIRMKIGVMFGNPEPTTGGNALKFYASVRLDIRRIGAVKEGENVVG	240
Qy	241	ETRVKVVKNKIAAPFKQAEFQILYGBGINFYGBELVDLGVKEKLEKAGAWYSYKGEKIQ	300
Db	241	ETRVKVVKNKIAAPFKQAEFQILYGBGINFYGBELVDLGVKEKLEKAGAWYSYKGEKIQ	300
Qy	301	GKANATAWLKNDPETAKEIEKKVRELLSNPNSTP	335
Db	301	GKANATAWLKNDPETAKEIEKKVRELLSNPNSTP	335

RESULT 4
US-10-733-782-122A-43319
; Sequence 43319, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangseu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 43319
; LENGTH: 333
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-282-122A-43319

Query Match 99.8%; Score 1674; DB 4; Length 353;
Best Local Similarity 99.7%; Pred. No. 7e-152;
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AIDENKOKALAAALGQIEKQFGKGSIMRLGEDRSMVKTISTGSLSDIALGAGGLPMGR 60
Db 2 AIDENKOKALAAALGQIEKQFGKGSIMRLGEDRSMVKTISTGSLSDIALGAGGLPMGR 61
Qy 61 IVEIYGPESSGKTTLTQVIAAAQREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 62 IVEIYGPESSGKTTLTQVIAAAQREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 121
Qy 121 TGEQALICDALARSGAVDVIVVDSVAALTPKABIEGIEGDSHMLAARMMWSQAWRKLAG 180
Db 122 TGEQALICDALARSGAVDVIVVDSVAALTPKABIEGIEGDSHMLAARMMWSQAWRKLAG 181
Qy 181 NLKQSNLTLLIFINQIRMKIGVMFNGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGVS 240
Db 182 NLKQSNLTLLIFINQIRMKIGVMFNGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGVS 241
Qy 241 ETRVKKVKNKIAAPFKQAEFQILLYGEGINFGYELVDLGKVEKLEKAGAWTSYKGEKIGQ 300
Db 242 ETRVKKVKNKIAAPFKQAEFQILLYGEGINFGYELVDLGKVEKLEKAGAWTSYKGEKIGQ 301
Qy 301 GKANATAWLNDKONPETAKEIEKKVRELLSNPNSTP 335
Db 302 GKANATAWLNDKONPETAKEIEKKVRELLSNPNSTP 336

RESULT 5

US-10-194-686-8
; Sequence 8, Application US/10194686
; Publication No. US20030148309A1
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIAN, VENKITESWATAN
; APPLICANT: CASTLE, LINDA A.
; APPLICANT: KREBER, CLAUD M.
; APPLICANT: BASS, STEVE
; APPLICANT: ZHANG, YING-XIN
; APPLICANT: COX, TONY
; APPLICANT: HUISMAN, GJALT
; APPLICANT: YUAN, LING
; APPLICANT: AFFHOLTER, JOSEPH A.
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
; FILE REFERENCE: 02-020730US
; CURRENT APPLICATION NUMBER: US/10/194,686
; CURRENT FILING DATE: 2002-07-11
; PRIOR APPLICATION NUMBER: US/09/354,922
; PRIOR FILING DATE: 1999-07-15
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-194-686-8

Query Match 99.8%; Score 1674; DB 4; Length 358;
Best Local Similarity 99.7%; Pred. No. 7.2e-152;
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AIDENKOKALAAALGQIEKQFGKGSIMRLGEDRSMVKTISTGSLSDIALGAGGLPMGR 60
Db 7 AIDENKOKALAAALGQIEKQFGKGSIMRLGEDRSMVKTISTGSLSDIALGAGGLPMGR 66

Qy 61 IVEIYGPESSGKTTLTQVIAAAQREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 67 IVEIYGPESSGKTTLTQVIAAAQREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 126
Qy 121 TGEQALICDALARSGAVDVIVVDSVAALTPKABIEGIEGDSHMLAARMMWSQAWRKLAG 180
Db 127 TGEQALICDALARSGAVDVIVVDSVAALTPKABIEGIEGDSHMLAARMMWSQAWRKLAG 186
Qy 181 NLKQSNLTLLIFINQIRMKIGVMFNGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGVS 240
Db 187 NLKQSNLTLLIFINQIRMKIGVMFNGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGVS 246
Qy 241 ETRVKKVKNKIAAPFKQAEFQILLYGEGINFGYELVDLGKVEKLEKAGAWTSYKGEKIGQ 300
Db 247 ETRVKKVKNKIAAPFKQAEFQILLYGEGINFGYELVDLGKVEKLEKAGAWTSYKGEKIGQ 306
Qy 301 GKANATAWLNDKONPETAKEIEKKVRELLSNPNSTP 335
Db 307 GKANATAWLNDKONPETAKEIEKKVRELLSNPNSTP 341

RESULT 6

US-10-194-686-10
; Sequence 10, Application US/10194686
; Publication No. US20030148309A1
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIAN, VENKITESWATAN
; APPLICANT: CASTLE, LINDA A.
; APPLICANT: KREBER, CLAUD M.
; APPLICANT: BASS, STEVE
; APPLICANT: ZHANG, YING-XIN
; APPLICANT: COX, TONY
; APPLICANT: HUISMAN, GJALT
; APPLICANT: YUAN, LING
; APPLICANT: AFFHOLTER, JOSEPH A.
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
; FILE REFERENCE: 02-020730US
; CURRENT APPLICATION NUMBER: US/10/194,686
; CURRENT FILING DATE: 2002-07-11
; PRIOR APPLICATION NUMBER: US/09/354,922
; PRIOR FILING DATE: 1999-07-15
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-194-686-10

Query Match 99.8%; Score 1674; DB 4; Length 358;
Best Local Similarity 99.7%; Pred. No. 7.2e-152;
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AIDENKOKALAAALGQIEKQFGKGSIMRLGEDRSMVKTISTGSLSDIALGAGGLPMGR 60
Db 7 AIDENKOKALAAALGQIEKQFGKGSIMRLGEDRSMVKTISTGSLSDIALGAGGLPMGR 66
Qy 61 IVEIYGPESSGKTTLTQVIAAAQREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 67 IVEIYGPESSGKTTLTQVIAAAQREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 126
Qy 121 TGEQALICDALARSGAVDVIVVDSVAALTPKABIEGIEGDSHMLAARMMWSQAWRKLAG 180
Db 127 TGEQALICDALARSGAVDVIVVDSVAALTPKABIEGIEGDSHMLAARMMWSQAWRKLAG 186
Qy 181 NLKQSNLTLLIFINQIRMKIGVMFNGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGVS 240

Db 187 NLKQSNLTLLIFINQIRMKIGVFNPGPTTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 246
Qy 241 ETRVKVVKNTAAAPKQAEFQILYEGEINFGYELVDLGKVKLEIKAGAWTSYKGEKIGQ 300
Db 247 ETRVKVVKNTAAAPKQAEFQILYEGEINFGYELVDLGKVKLEIKAGAWTSYKGEKIGQ 306
Qy 301 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 335
Db 307 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 341

RESULT 7
US-10-194-686-15
; Sequence 15, Application US/10194686
; Publication No. US20030148309A1
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIAN, VENKITESWATAN
; APPLICANT: CASTLE, LINDA A.
; APPLICANT: KREBBER, CLAUD M.
; APPLICANT: BASS, STEVE
; APPLICANT: ZHANG, YING-XIN
; APPLICANT: COX, TONY
; APPLICANT: HUISMAN, GJALT
; APPLICANT: YUAN, LING
; APPLICANT: AFFHOLTER, JOSEPH A.
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
; TITLE OF INVENTION: SEQUENCE RECOMBINATION
; FILE REFERENCE: 02-020730US
; CURRENT APPLICATION NUMBER: US/10/194,686
; PRIOR FILING DATE: 2002-07-11
; PRIOR APPLICATION NUMBER: US/09/354,922
; PRIOR FILING DATE: 1999-07-15
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 15
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: consensus
; OTHER INFORMATION: e. coli sequence
US-10-194-686-15

Query Match 99.8%; Score 1674; DB 4; Length 358;
Best Local Similarity 99.7%; Pred. No. 7.2e-152;
Matches 334; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AIDENKOKALAAALGQIEKQFGKGSIMRLGEDRSDMDVKTISTGSLSLDIALGAGGLPMGR 60
Db 7 AIDENKOKALAAALGQIEKQFGKGSIMRLGEDRSDMDVKTISTGSLSLDIALGAGGLPMGR 66
Qy 61 IVEIYGPSSGKTTTLQVIAAAQREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 67 IVEIYGPSSGKTTTLQVIAAAQREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 126
Qy 121 TGEQALEICDALARGAVDVIVDSVAALTPKAEIEGEGIDSHMGLAARMMSQAMRKLKAG 180
Db 127 TGEQALEICDALARGAVDVIVDSVAALTPKAEIEGEGIDSHMGLAARMMSQAMRKLKAG 186
Qy 181 NLKQSNLTLLIFINQIRMKIGVFNPGPTTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
Db 187 NLKQSNLTLLIFINQIRMKIGVFNPGPTTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 246
Qy 241 ETRVKVVKNTAAAPKQAEFQILYEGEINFGYELVDLGKVKLEIKAGAWTSYKGEKIGQ 300
Db 247 ETRVKVVKNTAAAPKQAEFQILYEGEINFGYELVDLGKVKLEIKAGAWTSYKGEKIGQ 306

Qy 301 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 335
Db 307 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 341

RESULT 8
US-10-194-686-9
; Sequence 9, Application US/10194686
; Publication No. US20030148309A1
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIAN, VENKITESWATAN
; APPLICANT: CASTLE, LINDA A.
; APPLICANT: KREBBER, CLAUD M.
; APPLICANT: BASS, STEVE
; APPLICANT: ZHANG, YING-XIN
; APPLICANT: COX, TONY
; APPLICANT: HUISMAN, GJALT
; APPLICANT: YUAN, LING
; APPLICANT: AFFHOLTER, JOSEPH A.
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
; TITLE OF INVENTION: SEQUENCE RECOMBINATION
; FILE REFERENCE: 02-020730US
; CURRENT APPLICATION NUMBER: US/10/194,686
; PRIOR FILING DATE: 2002-07-11
; PRIOR APPLICATION NUMBER: US/09/354,922
; PRIOR FILING DATE: 1999-07-15
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 9
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-194-686-9

Query Match 99.5%; Score 1669; DB 4; Length 358;
Best Local Similarity 99.1%; Pred. No. 2.2e-151;
Matches 332; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
Qy 1 AIDENKOKALAAALGQIEKQFGKGSIMRLGEDRSDMDVKTISTGSLSLDIALGAGGLPMGR 60
Db 7 AIDENKOKALATALGQIEKQFGKGSIMRLGEDRSDMDVKTISTGSLSLDIALGAGGLPMGR 66
Qy 61 IVEIYGPSSGKTTTLQVIAAAQREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 67 IVEIYGPSSGKTTTLQVIAAAQREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 126
Qy 121 TGEQALEICDALARGAVDVIVDSVAALTPKAEIEGEGIDSHMGLAARMMSQAMRKLKAG 180
Db 127 TGEQALEICDALARGAVDVIVDSVAALTPKAEIEGEGIDSHMGLAARMMSQAMRKLKAG 186
Qy 181 NLKQSNLTLLIFINQIRMKIGVFNPGPTTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
Db 187 NLKQSNLTLLIFINQIRMKIGVFNPGPTTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 246
Qy 241 ETRVKVVKNTAAAPKQAEFQILYEGEINFGYELVDLGKVKLEIKAGAWTSYKGEKIGQ 300
Db 247 ETRVKVVKNTAAAPKQAEFQILYEGEINFGYELVDLGKVKLEIKAGAWTSYKGEKIGQ 306
Qy 301 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 335
Db 307 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 341

RESULT 9
US-10-194-686-12
; Sequence 12, Application US/10194686

```
; Publication No. US20030148309A1
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIAN, VENKITESWATAN
; APPLICANT: CASTLE, LINDA A.
; APPLICANT: KREBBER, CLAUD M.
; APPLICANT: BASS, STEVE
; APPLICANT: ZHANG, YING-XIN
; APPLICANT: COX, TONY
; APPLICANT: HUISMAN, GJALT
; APPLICANT: AFFHOLTER, JOSEPH A.
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
; FILE REFERENCE: 02-020730US
; CURRENT APPLICATION NUMBER: US/10/194,686
; CURRENT FILING DATE: 2002-07-11
; PRIOR APPLICATION NUMBER: US/09/354,922
; PRIOR FILING DATE: 1999-07-15
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Escherichia coli
; US-10-194-686-12

Query Match          99.5%; Score 1669; DB 4; Length 358;
Best Local Similarity 99.1%; Pred. No. 2.2e-151;
Matches 332; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AIDENKOKALAAALGQIEKQFGKGSIMRLGEDRSMVKTISTGSLSDIALGAGGLPMGR 60
Db 7 AIDENKOKALAAALGQIEKQFGKGSIMRLGEDRSMVKTISTGSLSDIALGAGGLPMGR 66
Qy 61 IVEIYGPESSGKTTLTQVIAAAREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 67 IVEIYGPESSGKTTLTQVIAAAREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 126
Qy 121 TGEQALBICDALARSGAVDIVVDSVAALTPKABIEGIEGDSHMGGLAARMMSQAMRKLKAG 180
Db 127 TGEQALBICDALARSGAVDIVVDSVAALTPKABIEGIEGDSHMGGLAARMMSQAMRKLKAG 186
Qy 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
Db 187 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 246
Qy 241 ETRVKVVKNKIAAPFKQAEFQILYGEINFGYELVDLGVKEKLEKAGAWYSYKGEKIQ 300
Db 247 ETRVKVVKNKIAAPFKQAEFQILYGEINFGYELVDLGVKEKLEKAGAWYSYKGEKIQ 306
Qy 301 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 335
Db 307 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 341

RESULT 10
US-10-194-686-13
; Sequence 13, Application US/10194686
; Publication No. US20030148309A1
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIAN, VENKITESWATAN
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
```

```
; APPLICANT: CASTLE, LINDA A.
; APPLICANT: KREBBER, CLAUD M.
; APPLICANT: BASS, STEVE
; APPLICANT: ZHANG, YING-XIN
; APPLICANT: COX, TONY
; APPLICANT: HUISMAN, GJALT
; APPLICANT: YUAN, LING
; APPLICANT: AFFHOLTER, JOSEPH A.
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
; FILE REFERENCE: 02-020730US
; CURRENT APPLICATION NUMBER: US/10/194,686
; CURRENT FILING DATE: 2002-07-11
; PRIOR APPLICATION NUMBER: US/09/354,922
; PRIOR FILING DATE: 1999-07-15
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Escherichia coli
; US-10-194-686-13

Query Match          98.7%; Score 1656; DB 4; Length 358;
Best Local Similarity 98.5%; Pred. No. 3.8e-150;
Matches 330; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AIDENKOKALAAALGQIEKQFGKGSIMRLGEDRSMVKTISTGSLSDIALGAGGLPMGR 60
Db 7 AIDENKOKALAAALGQIEKQFGKGSIMRLGEDRSMVKTISTGSLSDIALGAGGLPMGR 66
Qy 61 IVEIYGPESSGKTTLTQVIAAAREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 120
Db 67 IVEIYGPESSGKTTLTQVIAAAREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD 126
Qy 121 TGEQALBICDALARSGAVDIVVDSVAALTPKABIEGIEGDSHMGGLAARMMSQAMRKLKAG 180
Db 127 TGEQALBICDALARSGAVDIVVDSVAALTPKABIEGIEGDSHMGGLAARMMSQAMRKLKAG 186
Qy 181 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 240
Db 187 NLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRIGAVKEGENVVGS 246
Qy 241 ETRVKVVKNKIAAPFKQAEFQILYGEINFGYELVDLGVKEKLEKAGAWYSYKGEKIQ 300
Db 247 ETRVKVVKNKIAAPFKQAEFQILYGEINFGYELVDLGVKEKLEKAGAWYSYKGEKIQ 306
Qy 301 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 335
Db 307 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 341

RESULT 11
US-10-194-686-11
; Sequence 11, Application US/10194686
; Publication No. US20030148309A1
; GENERAL INFORMATION:
; APPLICANT: DEL CARDAYRE, STEPHEN
; APPLICANT: TOBIN, MATTHEW
; APPLICANT: STEMMER, WILLEM P.C.
; APPLICANT: NESS, JON E.
; APPLICANT: MINSHULL, JEREMY
; APPLICANT: PATTEN, PHILLIP
; APPLICANT: SUBRAMANIAN, VENKITESWATAN
; APPLICANT: CASTLE, LINDA A.
; APPLICANT: KREBBER, CLAUD M.
; APPLICANT: ZHANG, YING-XIN
; APPLICANT: COX, TONY
; APPLICANT: HUISMAN, GJALT
; APPLICANT: YUAN, LING
; APPLICANT: AFFHOLTER, JOSEPH A.
; TITLE OF INVENTION: EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE
```

```
; TITLE OF INVENTION: SEQUENCE RECOMBINATION
; FILE REFERENCE: 02-020730US
; CURRENT APPLICATION NUMBER: US/10/194,686
; CURRENT FILING DATE: 2002-07-11
; PRIOR APPLICATION NUMBER: US/09/354,922
; PRIOR FILING DATE: 1999-07-15
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 358
; TYPE: PRN
; ORGANISM: Escherichia coli
US-10-194-686-11

Query Match      98.6%; Score 1655; DB 4; Length 358;
Best Local Similarity 98.8%; Pred. No. 4.8e-150;
Matches 331; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AIDENKOKALAAALGQIEKQFGKGSIMRLGDRSDMDVKTISTGSLSLDIALGAGGLPMGR 60
Db 7 AIDENKOKALAAALGQIEKQFGKGSIMRLGDRSDMDVKTISTGSLSLDIALGAGGLPMGR 66
Qy 61 IVEIYGPSSGKTTTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVVDIDNLLCSQPD 120
Db 67 IVEIYGPSSGKTTTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVVDIDNLLCSQPD 126
Qy 121 TGEQALEICDALARSGAVDIVVDSVAALTPKAEIEGIGDSHMGLAARMMSQAMRKLAG 180
Db 127 TGEQALEICDALARSGAVDIVVDSVAALTPKAEIEGIGDSHMGLAARMMSQAMRKLAG 186
Qy 181 NLKOSNTLLIFINQIRMKIGVMFGNPETTTGNNALKFYASVRLDIRRIGAVKEGENVVGS 240
Db 187 NLKOSNTLLIFINQIRMKIGVMFGNPETTTGNNALKFYASVRLDIRRIGAVKEGENVVGS 246
Qy 241 ETRVVKVKNKIAAPFKQAEFQILYEGEINFGYELVDLGVKEKLEKAGAWTSYKGEKIGQ 300
Db 247 ETRVVKVKNKIAAPFKQAEFQILYEGEINFGYELVDLGVKEKLEKAGAWTSYKGEKIGQ 306
Qy 301 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 335
Db 307 GKANAALWLGKGNPETAKEIEKKVRELLLSNPSTP 341

RESULT 12
US-10-282-122A-75948
; Sequence 75948, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELTRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347

; TITLE OF INVENTION: SEQUENCE RECOMBINATION
; FILE REFERENCE: 02-020730US
; CURRENT APPLICATION NUMBER: US/10/194,686
; CURRENT FILING DATE: 2002-07-11
; PRIOR APPLICATION NUMBER: US/09/354,922
; PRIOR FILING DATE: 1999-07-15
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 358
; TYPE: PRN
; ORGANISM: Salmonella typhi
US-10-282-122A-75948

Query Match      97.6%; Score 1637; DB 4; Length 353;
Best Local Similarity 97.0%; Pred. No. 2.5e-148;
Matches 325; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

Qy 1 AIDENKOKALAAALGQIEKQFGKGSIMRLGDRSDMDVKTISTGSLSLDIALGAGGLPMGR 60
Db 2 AIDENKOKALAAALGQIEKQFGKGSIMRLGDRSDMDVKTISTGSLSLDIALGAGGLPMGR 61
Qy 61 IVEIYGPSSGKTTTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVVDIDNLLCSQPD 120
Db 62 IVEIYGPSSGKTTTLQVIAAAQREGKTCFIDAEHALDPIYARKLGVVDIDNLLCSQPD 121
Qy 121 TGEQALEICDALARSGAVDIVVDSVAALTPKAEIEGIGDSHMGLAARMMSQAMRKLAG 180
Db 122 TGEQALEICDALARSGAVDIVVDSVAALTPKAEIEGIGDSHMGLAARMMSQAMRKLAG 181
Qy 181 NLKOSNTLLIFINQIRMKIGVMFGNPETTTGNNALKFYASVRLDIRRIGAVKEGENVVGS 240
Db 182 NLKOSNTLLIFINQIRMKIGVMFGNPETTTGNNALKFYASVRLDIRRIGAVKEGENVVGS 241
Qy 241 ETRVVKVKNKIAAPFKQAEFQILYEGEINFGYELVDLGVKEKLEKAGAWTSYKGEKIGQ 300
Db 242 ETRVVKVKNKIAAPFKQAEFQILYEGEINFGYELVDLGVKEKLEKAGAWTSYKGEKIGQ 301
Qy 301 GKANATAWLKONPETAKEIEKKVRELLLSNPSTP 335
Db 302 GKANATTWLNKPNATAKEIEKRVRELLLSNQNPATP 336

RESULT 13
US-10-282-122A-59525
; Sequence 59525, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELTRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
```


FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
PRIOR FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 59525
LENGTH: 352
TYPE: PRT
ORGANISM: Klebsiella pneumoniae
US-10-282-122A-59525

Query Match 95.9%; Score 1609; DB 4; Length 352;
Best Local Similarity 95.2%; Pred. No. 1.2e-145;
Matches 319; Conservative 9; Mismatches 7; Indels 0; Gaps 0;

Qy 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGSDRSMVKTISTGSLSLDIALGAGGLPMGR 60
Db 2 AIDENKQKALAAALGQIEKQFGKGSIMRLGSDRSMVKTISTGSLSLDIALGAGGLPMGR 61
Qy 61 IVEYGPSSGKTTTLQVIAAAREGKTCFIDAEHALDPYARKLGVDINDLLCSQPD 120
Db 62 IVEYGPSSGKTTTLQVIAAAREGKTCFIDAEHALDPYARKLGVDINDLLCSQPD 121
Qy 121 TGEALICDALARSGADVIVDVSVAALTPKAEIEGEGISHMGLAARMMSQAWRKLKAG 180
Db 122 TGEALICDALARSGADVIVDVSVAALTPKAEIEGEGISHMGLAARMMSQAWRKLKAG 181
Qy 181 NLQKSNLLIFINQIRMKIGVMFGNPTTTGGNALKFYASVRLDIRRIGAVKEGENVVGVS 240
Db 182 NLQKSNLLIFINQIRMKIGVMFGNPTTTGGNALKFYASVRLDIRRIGAVKEGENVVGVS 241
Qy 241 ETRVKVVKNTIAAPFKQAEFQILYGEINFGYELVDLGVEKLEKAGAWTSYKGEKIGQ 300
Db 242 ETRVKVVKNTIAAPFKQAEFQILYGEINFGYELVDLGVEKLEKAGAWTSYKGEKIGQ 301
Qy 301 GKANATAWLKDNPTAKEIEKKVRELLLSNPSTP 335
Db 302 GKANATWLKENPAAAKEIEKKVRELLLNQDQKP 336

RESULT 14
US-10-282-122A-55632
Sequence 55632, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
PRIOR FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 55632
LENGTH: 352
TYPE: PRT
ORGANISM: Enterobacter cloacae
US-10-282-122A-55632

Query Match 94.7%; Score 1589; DB 4; Length 352;
Best Local Similarity 94.3%; Pred. No. 1e-143;
Matches 316; Conservative 10; Mismatches 9; Indels 0; Gaps 0;

Qy 1 AIDENKQKALAAALGQIEKQFGKGSIMRLGSDRSMVKTISTGSLSLDIALGAGGLPMGR 60
Db 2 AIDENKQKALAAALGQIEKQFGKGSIMRLGSDRSMVKTISTGSLSLDIALGAGGLPMGR 61
Qy 61 IVEYGPSSGKTTTLQVIAAAREGKTCFIDAEHALDPYARKLGVDINDLLCSQPD 120
Db 62 IVEYGPSSGKTTTLQVIAAAREGKTCFIDAEHALDPYARKLGVDINDLLCSQPD 121
Qy 121 TGEALICDALARSGADVIVDVSVAALTPKAEIEGEGISHMGLAARMMSQAWRKLKAG 180
Db 122 TGEALICDALARSGADVIVDVSVAALTPKAEIEGEGISHMGLAARMMSQAWRKLKAG 181
Qy 181 NLQKSNLLIFINQIRMKIGVMFGNPTTTGGNALKFYASVRLDIRRIGAVKEGENVVGVS 240
Db 182 NLQKSNLLIFINQIRMKIGVMFGNPTTTGGNALKFYASVRLDIRRIGAVKEGENVVGVS 241
Qy 241 ETRVKVVKNTIAAPFKQAEFQILYGEINFGYELVDLGVEKLEKAGAWTSYKGEKIGQ 300
Db 242 ETRVKVVKNTIAAPFKQAEFQILYGEINFGYELVDLGVEKLEKAGAWTSYKGEKIGQ 301
Qy 301 GKANATAWLKDNPTAKEIEKKVRELLLSNPSTP 335
Db 302 GKANATWLKENPAAAKEIEKKVRELLLNQDQKP 336

RESULT 15
US-10-282-122A-78524
Sequence 78524, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 78524
; LENGTH: 356
; TYPE: PRT
; ORGANISM: Yersinia pestis
; US-10-282-122A-78524

Query Match 91.2%; Score 1530; DB 4; Length 356;
Best Local Similarity 91.6%; Pred. No. 4.8e-138;
Matches 304; Conservative 15; Mismatches 13; Indels 0; Gaps 0;

Qy	1	AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMDVKTISTGSLSLDIALGAGGLPMGR	60
Db	2	AIDENKQKALAAALGQIEKQFGKGSIMRLGEDRSMDVKTISTGSLSLDIALGAGGLPMGR	61
Qy	61	IVEIYGPSSGKTTTLTQVIAAAQREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD	120
Db	62	IVEIYGPSSGKTTTLTQVIAAAQREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPD	121
Qy	121	TGEQALEICDALARSGAVDVIIVDSVAALTPKAEIEGIGDSHMGSLAARMMSQAMRKLKAG	180
Db	122	TGEQALEICDALTSGAVDVIIVDSVAALTPKAEIEGIGDSHMGSLAARMMSQAMRKLKAG	181
Qy	181	NLKQSNFTLLIFINQIRMKIGVMFGNPETTTGKNALKFYASVRLDIRRIGAVKEGENVYGS	240
Db	182	NLKNANTLLIFINQIRMKIGVMFGNPETTTGKNALKFYASVRLDIRRIGAVKGDVVYGS	241
Qy	241	ETRVKVNKNKAAPFKQAEFQILYEGEINFGYELVDLGKKEKLEKAGAWTSYKGEKIGQ	300
Db	242	ETRVKVNKNKAAPFKQAEFQILYEGEINFGYELVDLGKKEKLEKAGAWTSYNGDKIGQ	301
Qy	301	GKANATAWLKONPETAKEIEKKVRELLSNPN	332
Db	302	GKANASNVLKENPAIAAELDKKREMLNGEN	333

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 16, 2006, 01:40:35 ; Search time 10.5 Seconds
(without alignments)
453.421 Million cell updates/sec

Title: US-10-733-782-3

Perfect score: 1678

Sequence: 1 AIDENKOKALAAALGQIEKQ.....AKEIEKKVRELLSNPNSTP 335

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 107799 seqs, 14211699 residues

Total number of hits satisfying chosen parameters: 107799

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA_New*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1208.5	72.0	375	6	US-10-467-657-4766 Sequence 4766, Ap
2	1019	60.7	363	7	US-11-074-176-296 Sequence 296, App
3	970.5	57.8	355	7	US-11-098-686-11427 Sequence 11427, A
4	154	9.2	443	7	US-11-098-686-10861 Sequence 10861, A
5	111.5	6.6	459	6	US-10-467-657-3092 Sequence 3092, App
6	109	6.5	453	6	US-10-467-657-206 Sequence 206, App
7	109	6.5	453	6	US-10-467-657-6400 Sequence 6400, Ap
8	104	6.2	454	7	US-11-098-686-10599 Sequence 10599, A
9	102	6.1	453	6	US-10-467-657-3626 Sequence 3626, Ap
10	100	6.0	925	6	US-10-454-437-50 Sequence 50, Appl
11	99	5.9	291	6	US-10-467-657-4148 Sequence 4148, Ap
12	97	5.8	376	7	US-11-213-368-14 Sequence 14, Appl
13	96	5.7	286	7	US-11-082-389-88 Sequence 88, Appl
14	96	5.7	471	6	US-10-467-657-802 Sequence 802, App
15	95	5.7	257	7	US-11-156-084-264 Sequence 264, App
16	95	5.7	376	7	US-11-082-389-122 Sequence 122, App
17	95	5.7	376	7	US-11-082-389-124 Sequence 124, App
18	94.5	5.6	548	6	US-10-793-626-326 Sequence 326, App
19	94.5	5.6	640	7	US-11-205-109-9 Sequence 9, Appl
20	94.5	5.6	716	6	US-10-131-826A-96 Sequence 96, Appl
21	94	5.6	365	7	US-11-082-389-132 Sequence 132, App
22	94	5.6	405	6	US-10-467-657-7420 Sequence 7420, Ap
23	94	5.6	479	7	US-11-082-389-130 Sequence 130, App
24	94	5.6	655	7	US-11-094-586-10 Sequence 10, Appl
25	92.5	5.5	423	7	US-11-070-080-16 Sequence 16, Appl

26	92	5.5	302	6	US-10-793-626-2798 Sequence 2798, Ap
27	92	5.5	376	7	US-11-213-368-6 Sequence 6, Appli
28	92	5.5	650	6	US-10-878-556A-110 Sequence 110, App
29	92	5.5	806	7	US-11-108-172-1117 Sequence 1117, Ap
30	92	5.5	806	7	US-11-108-172-1118 Sequence 1118, Ap
31	92	5.5	806	7	US-11-079-900-1 Sequence 1, Appli
32	91.5	5.5	419	6	US-10-979-821-6 Sequence 6, Appli
33	91.5	5.5	419	7	US-11-114-922-6 Sequence 6, Appli
34	91.5	5.5	569	6	US-10-506-443A-35 Sequence 35, Appl
35	90.5	5.4	718	7	US-11-074-176-306 Sequence 306, App
36	90.5	5.4	723	7	US-11-074-176-18 Sequence 18, Appl
37	89.5	5.3	690	7	US-11-212-443-171 Sequence 171, App
38	89.5	5.3	710	7	US-11-212-443-30 Sequence 30, Appl
39	89	5.3	885	6	US-10-467-657-2302 Sequence 2302, Ap
40	89	5.3	1992	7	US-11-013-759-3 Sequence 3, Appli
41	89	5.3	1992	7	US-11-013-759-13 Sequence 13, Appl
42	89	5.3	2047	7	US-11-013-759-4 Sequence 4, Appli
43	89	5.3	2047	7	US-11-013-759-7 Sequence 7, Appli
44	88.5	5.3	687	7	US-11-212-443-173 Sequence 173, App
45	88.5	5.3	710	7	US-11-212-443-28 Sequence 28, Appl

ALIGNMENTS

RESULT 1

US-10-467-657-4766
; Sequence 4766, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10467.657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 4766
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-4766

Query Match		72.0%	Score 1208.5;	DB 6;	Length 375;
Best Local Similarity		68.7%	Pred. No. 5.8e-86;		
Matches 224;		Conservative 53;	Mismatches 48;	Indels 1;	Gaps 1;
Qy	2	IDENKOKALAAALGQIEKQFGKGSIMRL-GEDRSMVKTITSTGSLSLDIALGAGGLPMGR	60		
Db	28	MSDDKSKALAAALQIEKSGKAIMQDGSQBEENLEVIETGSLGLDLALGVGGLPRGR	87		
Qy	61	IVEIYGESSGKTTLTLOVIAAARECKTCAPDAEHALDPIYARKLGVDIDNLLCSQPD	120		
Db	88	IVEIFGESSGKTTLCLEAVACQKNGGVCAFDVAEHAFFVYARKLGKVBELYSQPD	147		
Qy	121	TGEQALIEICALARSGAVDVIVVDSVAALTPKABIEGIEIGDSHGLAARMMQAMRKLKAG	180		
Db	148	TGEQALIEICUTLVRSGGIDVWVDSVAALVPKABIEGDMGDSHVGLQARLMSQALRKLTG	207		
Qy	181	NLKQSNLLIFINQIRMKIGVMFENPETTTGGNALKFPYASVRLDIRRIGAVKEGENVVGVS	240		
Db	208	HIKKTNTLVVFINQIRMKIGVMFGSPETTTGGNALKFPYSSVRLDIRRTGSIKKGEVLGN	267		
Qy	241	ETRVKVKNKIAAPFKQAEFOILYGGGINFYGVLDLGVKEKLEKAGAWSYKGEKIGQ	300		
Db	268	ETRVKVKINKVAPPFQAEFDILYGEISWEGELIDITGVKNNDIINKSGAWSYNGAKIGQ	327		

```
Qy 301 GKANATAWLKNDNPETAKEIEKKVREL 326
Db 328 GKDNVRVWLKSNPEISDAIKRAL 353

RESULT 2
US-11-074-176-296
; Sequence 296, Application US/11074176
; Publication No. US20050250135A1
; GENERAL INFORMATION:
; APPLICANT: Klaenhammer, Todd R.
; APPLICANT: Russell, William M.
; APPLICANT: Altermann, Eric
; APPLICANT: McAuliffe, Olivia
; APPLICANT: Perill, Andrea Azcarate
; TITLE OF INVENTION: Nucleic Acid Sequences Encoding
; TITLE OF INVENTION: Stress-Related Proteins and Uses Therefore
; FILE REFERENCE: 5051-694
; CURRENT APPLICATION NUMBER: US/11/074,176
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: 60/551,161
; PRIOR FILING DATE: 2004-03-08
; NUMBER OF SEQ ID NOS: 381
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 296
; LENGTH: 363
; TYPE: PRT
; ORGANISM: Lactobacillus acidophilus
US-11-074-176-296

Query Match 60.7%; Score 1019; DB 7; Length 363;
Best Local Similarity 56.8%; Pred. No. 2.2e-71;
Matches 183; Conservative 74; Mismatches 65; Indels 0; Gaps 0;

Qy 4 ENKOKALAAALGOIEKQFGKSGIMRLGEDRSMDVKTISTGSLSDIALGAGGLPMGRIVE 63
Db 4 DEKKAALDRAALKKIEKNGFGKAVMRGKEKADTQISTVTGSLDAAIGVGYPGRILIE 63

Qy 64 IYGPSSGKTTLTLOVIAAAQREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPDGTGE 123
Db 64 VYGPSSGKTTVALHAVAQVKGRTAAVYDAENAMDPAAYEAENGLVDIDSLISQPDGTGE 123

Qy 124 QALEICDALARSGAVDVIVVDSVAALTPKAEIEGIDSHMGLAARMMSQAMRKLGNLX 183
Db 124 EGLQADTLTSSGADIIVVDSVAALVPRAEIEGEMGDHVGLOQLRLMSQALRLKLSGTIS 183

Qy 184 OSNTLLIFINQIRMKIGVMFGNPETTTGNNALKFYASVRLDIRRIGAVKEGENVVGSETR 243
Db 184 KTKTIAIFINQIREKVGVMFGNPETTPGGRALKFYSTVRLVRRAEQIKQSGDVLGNRVK 243

Qy 244 VVKVKNKIAAPPKQAEFOILYGEINFGYELVDLGKVEKLEKAGAWSYKGEKIQGKA 303
Db 244 IKVVKNKVAPPKVAEVDIMYKGISQSGELLDMAADKDIIDKAGSWYSYKSDRIGQGRE 303

Qy 304 NATAWLKNDNPETAKEIEKKVRE 325
Db 304 NAKKYLEBHPDIYQVQSVQRQ 325

RESULT 3
US-11-098-686-11427
; Sequence 11427, Application US/11098686
; Publication No. US20060024696A1
; GENERAL INFORMATION:
; APPLICANT: Kapur, Vivek and Gebhart, Connie J.
; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES
; TITLE OF INVENTION: FROM LAWSONIA INTRACELLULARIS AND METHODS OF USING
; FILE REFERENCE: 09531-128001
; CURRENT APPLICATION NUMBER: US/11/098,686
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: PCT/US03/31318
; PRIOR FILING DATE: 2003-10-01
; NUMBER OF SEQ ID NOS: 11433
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10861
; LENGTH: 443
; TYPE: PRT
; ORGANISM: Lawsonia intracellularis
US-11-098-686-11427

Query Match 9.2%; Score 154; DB 7; Length 443;
Best Local Similarity 32.1%; Pred. No. 0.00011;
Matches 62; Conservative 28; Mismatches 81; Indels 22; Gaps 8;

Qy 18 EKQFGKSGIMRLGEDRSMDVKTISTGSLSDIALGAGGLPMGRIVEIYGPSSGKTTTL 77
Db 44 QKTYKKLPIDPIGRVETSTYQPFSTGKVLHDILKGLVPSALL-IGEGIGIGKSTLL 102

Qy 78 QVIAAAREGKTCAFIDAEHALDPIYAR--KLGVVDIDNLLCSQPDGTGEQALEICDALAR 135
Db 103 QLQATATLGLKVLVYVSGEESLPQIKARAERLTVLHNDLLSM--ATSQVEDILPLNTS 159
```

```
; PRIOR FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 11433
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11427
; LENGTH: 355
; TYPE: PRT
; ORGANISM: Lawsonia intracellularis
US-11-098-686-11427

Query Match 57.8%; Score 970.5; DB 7; Length 355;
Best Local Similarity 57.3%; Pred. No. 1.1e-67;
Matches 189; Conservative 64; Mismatches 76; Indels 1; Gaps 1;

Qy 3 DENKOKALAAALGOIEKQFGKSGIMRLGEDRSMDVKTISTGSLSDIALGAGGLPMGRIV 62
Db 9 EDSRREALKTALDTIERKFGQGVKLSDDVVHVAVIPTSGISGLDLAGLGGIPRGV 68

Qy 63 EIVGPSSGKTTLTLOVIAAAQREGKTCAFIDAEHALDPIYARKLGVDIDNLLCSQPDGT 122
Db 69 EIVGPSSGKTTLTHTIAECQKLGTTAAFDIAEHALDIAVAKRLGVKTDLEIISQPDHG 128

Qy 123 EQALEICDALARSGAVDVIVVDSVAALTPKAEIEGIDSHMGLAARMMSQAMRKLGNL 182
Db 129 EQALEIADMLVRSQAVDLVVVDSVAALIPOTELESGMGETQVGGHARLMSHALRKLGTI 188

Qy 183 KQNTLLIFINQIRMKIGVM-FGNPETTTGNNALKFYASVRLDIRRIGAVKEGENVVSE 241
Db 189 HKSHTAVIFINQIRMKIGVVGYSPTTTGNNALKFYSSVNRMDIRKIOTLKDKBESYSSL 248

Qy 242 TRVKVKNKIAAPPKQAEFOILYGEINFGYELVDLGKVEKLEKAGAWSYKGEKIQGQ 301
Db 249 TRVKVKNKVAPPKPAKFDIINGTISRSGELIDLGVEAGIIDKSGSWFAFGSEKLGQ 308

Qy 302 KANATAWLKNDNPETAKEIEKKVRELLSNP 331
Db 309 KEKVRALLDENTSLKSSIETALMEHLGNP 338
```

```
RESULT 4
US-11-098-686-10861
; Sequence 10861, Application US/11098686
; Publication No. US20060024696A1
; GENERAL INFORMATION:
; APPLICANT: Kapur, Vivek and Gebhart, Connie J.
; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES
; TITLE OF INVENTION: FROM LAWSONIA INTRACELLULARIS AND METHODS OF USING
; FILE REFERENCE: 09531-128001
; CURRENT APPLICATION NUMBER: US/11/098,686
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: PCT/US03/31318
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: US 60/416,395
; PRIOR FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 11433
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10861
; LENGTH: 443
; TYPE: PRT
; ORGANISM: Lawsonia intracellularis
US-11-098-686-10861
```

```
Query Match 9.2%; Score 154; DB 7; Length 443;
Best Local Similarity 32.1%; Pred. No. 0.00011;
Matches 62; Conservative 28; Mismatches 81; Indels 22; Gaps 8;

Qy 18 EKQFGKSGIMRLGEDRSMDVKTISTGSLSDIALGAGGLPMGRIVEIYGPSSGKTTTL 77
Db 44 QKTYKKLPIDPIGRVETSTYQPFSTGKVLHDILKGLVPSALL-IGEGIGIGKSTLL 102

Qy 78 QVIAAAREGKTCAFIDAEHALDPIYAR--KLGVVDIDNLLCSQPDGTGEQALEICDALAR 135
Db 103 QLQATATLGLKVLVYVSGEESLPQIKARAERLTVLHNDLLSM--ATSQVEDILPLNTS 159
```


Qy 147 AALTPKA--EIEGEIGDSHMGAAARMMSQAMRKLAGNLKQSNLTLLIFINQIRMKIGVMFG 204
Db 321 -HIMPRAGRDEVAELGN-----ISRRLKNLAEL---NTPVVLVAQLNR-----G 361
Qy 205 NPETTTGNNALKFYASVR---LDIRIRIGAVKEGENVV 238
Db 362 NTK-----QADKEPNWADIRGSGAIEQDANII 388

RESULT 8

US-11-098-686-10599
; Sequence 10599, Application US/11098686
; Publication No. US20060024696A1
; GENERAL INFORMATION:
; APPLICANT: Kapur, Vivek and Gebhart, Connie J.
; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES
; FILE OF INVENTION: FROM LAWSONIA INTRACELLULARIS AND METHODS OF USING
; FILE REFERENCE: 09531-128001
; CURRENT APPLICATION NUMBER: US/11/098,686
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: PCT/US03/31318
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: US 60/416,395
; PRIOR FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 11433
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10599
; LENGTH: 454
; TYPE: PRT
; ORGANISM: Lawsonia intracellularis
US-11-098-686-10599

Query Match 6.2%; Score 104; DB 7; Length 454;

Best Local Similarity 22.0%; Pred. No. 0.84;
Matches 55; Conservative 45; Mismatches 94; Indels 56; Gaps 11;

Qy 16 QIEQFGKSGIMRLGEDRSDMDVKITSGSLDIALGAGGLPMGRIVEIYGPSSGKTYL 75
Db 168 ELEKRF-----ERKEQVTGTTGYNRLD-KLTAGLQFSDLIIVAAAP-SMGKTAF 215
Qy 76 TLQV-IAAAREGKTCAPIDAEHALDPIYAKL-----GVIDNLLCS--QPTDGEQALEI 128
Db 216 SLNWMRSAIQGTPTVAIYSLEMSNNQLMWMCLCWGKVDLSLHRLHGYLNSDEWSRLYHA 275
Qy 129 CDALARSQAVDIVVDSVAALTP-----KAEIEGEI-----GDSHMGAA 168
Db 276 ADVLGQA---PIFIDTTPALSPLELRARTRRLKIESDVGLWIDYQLMRGNKRTDSRE 331
Qy 169 RMMSQAMRKLAGNLKQSNLTLLIFINQIRMKIGVMFGNPETTTGGNNALKFYASVRLDIRRI 228
Db 332 QEISISRLSKSLAKEINIPVVALSQLNRKL-----EDRTDKRPQLS-----DURES 378
Qy 229 GAVKEGENVV 238
Db 379 GAIEQDADVI 388

RESULT 9

US-10-467-657-3626
; Sequence 3626, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8

; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 3626
; LENGTH: 453
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-3626

Query Match 6.1%; Score 102; DB 6; Length 453;

Best Local Similarity 23.7%; Pred. No. 1.2;
Matches 69; Conservative 43; Mismatches 85; Indels 94; Gaps 18;

Qy 2 IDENKQKA--LAAALQIEKQFGKGSIMRL-----GKDAVKRETTFGQTVEDLIGLDRLOGVRFGLPTG 43
Db 138 VAEKLSRAADELAAA-----GKDAVKRETTFGQTVEDLIGLDRLOGVRFGLPTG 189
Qy 44 SLSLDIALG-AGGLPMGRIVEIYGPSSGKTYLTLQVIAAAQREGKTCAPIDAEHALDPI 102
Db 190 LMKLD---GMTGGLPDGNLIVIAARPSMGKTVLAENIARFALKQKQKAVHFQSYEMSAVEL 246
Qy 103 ----YARKLGVDIDNL-----LCSQPDTEQALE-----LCDAL----- 132
Db 247 ARRGMAAECNIPMQLNLTQSDYANMPITYVSOAKWKFPDVCNCDLLNVDELCLFLAKEK 306
Qy 133 ARSGAVDIVVDSVAALTPKA--EIEGEIGDSHMGAAARMMSQAMRKLAGNLKQSNLTLLI 190
Db 307 KLTTGLDLLVVDHL-HIMPRAGRDEVAELGN-----ISRRLKNLAEL---NTPVV 353
Qy 191 FINQIRMKIGVMFGNPETTTGGNNALKFYASVR---LDIRIRIGAVKEGENVV 238
Db 354 LVAQLNR-----GNTK-----QADKEPNWADIRGSGAIEQDANII 388

RESULT 10

US-10-454-437-50
; Sequence 50, Application US/10454437
; Publication No. US20050277115A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; FILE REFERENCE: BGI-128CPCN
; CURRENT APPLICATION NUMBER: US/10/454,437
; CURRENT FILING DATE: 2003-06-13
; PRIOR APPLICATION NUMBER: US 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: DE 199331636.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 199332125.6
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 199332126.4
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 199332127.2
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 199332128.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 199332129.9
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: DE 199332226.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 199332920.6
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 199332922.2
; PRIOR FILING DATE: 1999-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 442
; SEQ ID NO 50
; LENGTH: 925

```
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-10-454-437-50

Query Match
Best Local Similarity 6.0%; Score 100; DB 6; Length 925;
Matches 75; Conservative 67; Mismatches 125; Indels 120; Gaps 18;

Qy 14 LQIEKQFGKGS--IMRLGED-----RSMVDKTIIS-----TGSLSLDIALGAG 54
Db 110 LGLIREGEGVAAQVLVLKGLADLPVRVQVQVQLLSGVEGGSPGGQGAPGGDAVCGAG 169

Qy 55 GLPWGRIVEIIVGPSSG-----KTYLTL-----QVTAARQREGKTCFAIDAEHALDPIYAR 105
Db 170 AAPGGR-----PSSGSGFGERSTSLVLDQFGRNLTOAKDGK-----LDPVVGR 212

Qy 106 KLGVD-----IDNLCSQPDTCGEQALEICDALARSGAVDVIIVDSVAALTPKAE 154
Db 213 DKEIERIMQVLSRRTKNNPVLIGEPGVGTAV-----VEGLALDIVNGKVPETL 261

Qy 155 IEGEIGDSHMG--LAARM-----MSQAMKRLAGNLKQSNLTLLIFINQIRMKIGVMFGNPET 208
Db 262 KDKQVYSLDLGSLVAGSYRGDPFERLKKVLKEINQRGDIILFIDEIHTLVGA--GAERG 319

Qy 209 TTGGNALKFYASVRLDIRRIGAV-----KGENVVGSETRVKVVKNKIAAPFKQAEFQI 262
Db 320 AIDAASLLKPKLARGELQTIGATTILDEYRKHIEKDAALERFPQV--QVPEFSVDLTVEI 377

Qy 263 LYG-----EGINFYGEVLVGLVKEK-----LIEKAGANYSYKGEKIG 299
Db 378 LKGURDREYEAHRVSIITGATAAQAQADRVINDRFLPDKAVDLIDEAGARMRIK----- 432

Qy 300 QGKANATAWLNDKNPETAKEIKKVVREL 326
Db 433 -----RWTA-----PSSLREVDERTADV 450

RESULT 11
US-10-467-657-4148
; Sequence 4148, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 4148
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-4148

Query Match
Best Local Similarity 5.9%; Score 99; DB 6; Length 291;
Matches 63; Conservative 44; Mismatches 83; Indels 76; Gaps 15;

Qy 55 GLPW---GRIVEIIVGPSSGKTTTLQVIAAARQREGKTCFAID--AEHALDPIYARKLV 109
Db 70 GIPLNNRQKVIAIDPGGASGKGTVAARKVAALGYD-----YLDTGALYRLTALYARQGV 124

Qy 110 ---DIDNLCSQPDTCGEQALEICDALAR-----SGAVDVIIVDSVAALTPKABIEGIG 160
Db 125 EWHDEENV-----SALAKGLPAVFSG--NRILLDG-----EDVSDGIR 160

; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-10-454-437-50

Query Match
Best Local Similarity 5.8%; Score 97; DB 7; Length 376;
Matches 53; Conservative 30; Mismatches 80; Indels 62; Gaps 11;

Qy 58 MGRIVEIY--GPSSSGKTTTLQVIAAARQREGKTCFAIDAEHALDPI-YARKLVGD--ID 112
Db 46 MPILLRVYIDGPHGKGTITTLQVIALGSRD-----DIVYVDPPTWQVLAGSETIA 98

Qy 113 NLCSQP-----DTGEQALEICDALARSGAVDVIIVDSVAALTPKABIEGIGDSHM- 164
Db 99 NIVTQHRLDQGEISAGDAAVVMTSAQITWGM-PPYAVTDAVLA-----PHIGGAGSSHAP 153

Qy 165 -----GLAARMMSQAMKRLAGNLKQSNLTLLIFINQIRMKIGVMFGNPETTTGCGN 213
Db 154 PPALTILFDRHRPIAALCYPAARYLMGSM--TPQAVLAFVALI-----PPTLPCTN 202

Qy 214 -----ALKFYASVRLDIRRIGAVKGENVVGSETR 243
Db 203 IVIGALPEDRHIDRLAKRQRPGERLDLAMLAIIRRVYGLLANTVR 247

RESULT 13
US-11-082-389-88
; Sequence 88, Application US/11082389
; Publication No. US20050244935A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; TITLE OF INVENTION: CORNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; TITLE OF INVENTION: INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE
; TITLE OF INVENTION: TRANSPORT
```


Db 196 YFANQYDLSEVETLIKRDSSRRYAKKQLTYFKNQFPDMRWEDTKQITENPKLIIDLVKKFN 255
QY 325 E 325
Db 256 Q 256

Search completed: February 16, 2006, 01:45:02
Job time : 11.5 secs

This Page Blank (uspto)